

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE INFORMÁTICA



TRABAJO DE FIN DE GRADO

---

**ClassRoom VR-Motion Capture (CVR-MC):**

**un juego en RV para la mejora de la expresión corporal de los profesores de secundaria.**

ClassRoom VR-Motion Capture (CVR-MC): a VR game to improve corporal expression in secondary-school teachers.

---

Grado en Desarrollo de Videojuegos y Grado en Ingeniería del software

**Autores:**

**Sandra Alonso Paz**

**Daniel López Acero**

**Andrés Puente Rodríguez**

**Antonio Luis Suárez Solís**

**Directores:**

**Borja Manero Iglesias**

**Alejandro Romero Hernández**

**Curso académico 2020-2021**

# Agradecimientos:

Nos gustaría agradecer a todas las personas que nos han apoyado durante el proceso y que por tanto, han hecho posible que este proyecto se haya realizado con éxito.

En primer lugar, agradecemos a nuestros tutores Borja Manero y Alejandro Romero, por habernos dado la oportunidad de participar en este proyecto y contagiarnos su entusiasmo por el campo de la comunicación y la educación. Agradecer también su paciencia, dedicación y motivación en todo momento y que han hecho que hayamos podido mantener la ilusión y las ganas a lo largo de todo el año. De ellos no solo nos llevamos los conocimientos adquiridos, sino también una forma de entender el trabajo y la vida.

Por otra parte, nos gustaría agradecer toda la ayuda recibida por el equipo Didascalia y más concretamente a Ibis, Isabel, Montse y Miriela. A pesar de nuestros escasos conocimientos en psicología y pedagogía, supieron cómo contagiarnos su entusiasmo por sus campos de estudio y hacer que comprendiéramos todos los conceptos relacionados con ellos.

También nos gustaría destacar la colaboración recibida por Iván García-Magariño que nos permitió utilizar Emopose para el desarrollo de CVR-MC. Esta colaboración fue vital para el desarrollo de la aplicación.

Además agradecemos tanto a la Universitat Autònoma de Barcelona y a la Universidad Complutense de Madrid, y más concretamente a la Facultad de Informática, por habernos provisto de espacios adaptados para el desarrollo de nuestro proyecto.

Por último nos gustaría agradecer a nuestras familias y amigos que en todo momento han sido un apoyo moral indispensable para todos nosotros y que han sabido animarnos para superar los obstáculos que un proyecto de esta magnitud conlleva.

Gracias a todos.

# Resumen:

Contar con docentes cualificados y experimentados al frente de las aulas es imprescindible en cualquier sociedad. Sin embargo, la realidad es que no todos los docentes han podido disfrutar de una preparación práctica que les permita gestionar situaciones con las que no están familiarizados. Un ejemplo de estas, son las situaciones conflictivas. La falta de entornos seguros que permitan estas prácticas hace que se cree un problema insostenible al que debemos dar solución cuanto antes.

A la hora de afrontar un conflicto es realmente importante ser asertivo y mostrar corporalmente lo que queremos decir con nuestras palabras. La comunicación efectiva es una asignatura pendiente para la mayoría de docentes que comienzan su carrera profesional. Sin embargo, es algo que puede enseñarse y que mejora las habilidades de resolución de conflictos.

En este trabajo, presentamos la implementación y evaluación de Classroom VR-Motion Capture, un herramienta que permite a los docentes noveles situarse al frente de una clase donde se producen situaciones disruptivas y reaccionar ante ellas de manera segura. Durante la ejecución se recogerán datos relativos al lenguaje no verbal como son: los cambios en la tonalidad de la voz, los gestos y posiciones, la distancia entre interlocutores y palabras clave. Con esta información estimaremos la emoción con la que se relaciona el lenguaje no verbal del usuario. Además, tras su actuación recibirá un feedback valorando las decisiones tomadas para resolver el conflicto y un análisis del lenguaje no verbal empleado junto con las emociones estimadas.

Tras el desarrollo de la aplicación realizamos un experimento junto con 14 profesionales del sector de la educación de Barcelona. En este documento se describe la prueba realizada con Classroom VR-Motion Capture con la intención de responder a dos preguntas de investigación. ¿Es factible utilizar el sistema CVR-MC en la formación docente para contribuir al aprendizaje de competencias comunicativas de gestión de clima de aula? Por otra parte, ¿es posible capturar el lenguaje no verbal y las emociones que manifiestan los participantes durante la simulación? ¿Se corresponden con las que manifiestan en un entorno real?

De esta prueba concluimos que es un entorno amigable, seguro y factible para la preparación de futuros docentes. Sin embargo, hacemos una reflexión sobre cómo nuestro lenguaje no verbal y por lo tanto, las emociones que transmitimos, no concuerdan en entornos reales y entornos virtuales.

**Palabras clave:** Videojuegos, educación, realidad virtual, emociones, captura de movimiento, lenguaje no verbal.

# Abstract:

Nowadays having qualified and experienced teachers in school classrooms is considered to be of the highest priority in any society. Unlikely this is far from reality because most teachers confess, they don't have received enough practical training to manage disruptive situations in the classroom. The lack of safe environments which allow these practices creates an unsustainable problem which must be solved as soon as possible.

Facing conflict is usually a hard task and it is extremely important to be assertive and to show with your body what we want to say with our words. Effective communication seems to be an enormous issue for most teachers beginning their professional careers. Fortunately, it can be taught and also could improve conflict resolution skills.

In this project, we show not just the implementation but also the evaluation of Classroom VR-Motion Capture, an application which allows new teachers to experience different disruptive situations in a classroom and react to them in a safe environment. During the simulation, data related to non-verbal language will be collected, such as: voice tone variations, gestures and positions, distance between interlocutors (proxemia) and keywords. As a result of this procedure, we will estimate the emotion related to the user's non-verbal language. In addition, after their performance, they will receive feedback about the decisions made to resolve the conflict, an analysis of the non-verbal language used and the estimated emotions.

Coming up next to developing the application, we carried out an experiment where took part 14 education professionals from Barcelona. This document describes the test carried out with Classroom VR-Motion Capture for answering two research questions. Firstly, is it possible to use CVR-MC system in teacher training to improve the communications skills for classroom climate management? Secondly, is it possible to capture non-verbal language and their relevant emotions which are expressed by the participants during the simulation? Do they match with those expressed in real environments?

Thanks to this test we conclude that Classroom VR-Motion Capture is a friendly, safe and feasible environment for training future teachers. However, we observed that our non-verbal language and therefore the emotions we transmit, do not match in real and virtual environments.

**Keywords:** Video games, education, virtual reality, emotions, motion capture, non-verbal language.



# Índice general

<b>1</b>	<b>Introducción</b>	<b>11</b>
1.1.	Motivación . . . . .	11
1.2.	Objetivos . . . . .	12
1.2.1.	Objetivos generales . . . . .	12
1.2.2.	Objetivos específicos . . . . .	12
1.3.	Estructura del documento . . . . .	13
<b>2</b>	<b>Estado del arte</b>	<b>14</b>
2.1.	La comunicación . . . . .	14
2.1.1.	Comunicación no verbal . . . . .	15
2.1.1.1.	La importancia de la comunicación no verbal . . . . .	15
2.1.1.2.	Comunicación no verbal y las emociones . . . . .	16
2.2.	Docencia . . . . .	17
2.2.1.	Educación infantil y primaria . . . . .	17
2.2.2.	Educación secundaria . . . . .	17
2.3.	Gestión del aula . . . . .	18
2.3.1.	Gestión de conflictos . . . . .	18
2.3.1.1.	Qué es un conflicto . . . . .	18
2.3.1.2.	Formación del profesorado en gestión de conflictos . . . . .	19
2.3.2.	Importancia de aprender a gestionar el aula . . . . .	19
2.3.3.	Importancia de la comunicación en el manejo del aula . . . . .	19
2.4.	Tecnologías y comunicación no verbal . . . . .	20
2.4.1.	Herramientas que permiten la captura del lenguaje corporal . . . . .	20
2.4.1.1.	Kinect . . . . .	20
2.4.1.2.	Axis Neuron . . . . .	21
2.4.1.3.	Smartsuit Pro . . . . .	22
2.4.1.4.	OptiTrack . . . . .	22
2.4.1.5.	Xnet . . . . .	23
2.4.1.6.	Neuroelectrics Enobio 32 . . . . .	23
2.4.1.7.	ML Kit . . . . .	24
2.4.2.	Herramientas que permiten su análisis . . . . .	25
2.4.2.1.	Active Appearance models (AAM) . . . . .	25
2.4.2.2.	Redes Neuronales Convolucionales . . . . .	25
2.4.3.	Herramientas que extraen emociones del lenguaje corporal . . . . .	26
2.4.3.1.	Emopose . . . . .	26
2.4.3.2.	iMotions . . . . .	27
2.4.4.	Eye-Tracking Analysis for Emotions Recognition . . . . .	28

2.5.	Entornos automáticos de aprendizaje . . . . .	29
2.5.1.	Videojuegos educativos (serious games) . . . . .	29
2.5.1.1.	Qué son . . . . .	29
2.5.1.2.	Classroom VR . . . . .	29
2.5.1.3.	Didascalia (RTI2018-096401-A-I00) . . . . .	30
2.5.2.	La Realidad Virtual . . . . .	30
2.5.2.1.	Qué es . . . . .	30
2.5.2.2.	Aplicaciones de la Realidad Virtual . . . . .	31
2.5.2.3.	Realidad virtual y las habilidades de comunicación del profesorado . . . . .	33
2.5.2.4.	Ventajas e inconvenientes de la Realidad Virtual . . . . .	34
<b>3</b>	<b>Metodología y tecnología</b>	<b>36</b>
3.1.	Metodología de trabajo . . . . .	36
3.1.1.	Metodología empleada . . . . .	36
3.1.1.1.	Extreme programming (XP) . . . . .	37
3.1.2.	Equipo XP . . . . .	37
3.1.2.1.	Ciclo de vida XP . . . . .	38
3.1.3.	Planificación temporal . . . . .	38
3.1.4.	Control de versiones . . . . .	40
3.1.4.1.	GitHub . . . . .	40
3.1.4.2.	Google Drive . . . . .	41
3.2.	Tecnología base . . . . .	41
3.2.1.	Unity . . . . .	41
3.2.2.	C# . . . . .	42
3.3.	Software utilizado . . . . .	42
3.3.1.	Emopose . . . . .	42
3.3.2.	Axis Neuron . . . . .	42
3.3.3.	Software en pruebas . . . . .	44
3.3.3.1.	MLKit . . . . .	44
3.3.3.2.	ThreeDPoseUnityBarracuda . . . . .	44
3.4.	Dispositivos utilizados . . . . .	45
3.4.1.	Oculus Rift S . . . . .	45
3.4.2.	Perception Neuron . . . . .	46
3.4.3.	Ordenador para ejecutar el programa . . . . .	47
3.4.4.	Dispositivos en fase de pruebas . . . . .	47
3.4.4.1.	Kinect . . . . .	47
3.4.4.2.	Gafas Emteq Labs . . . . .	47
<b>4</b>	<b>Diseño y desarrollo de la herramienta</b>	<b>49</b>
4.1.	Diseño del proyecto . . . . .	49
4.1.1.	Fase divergente . . . . .	49
4.1.2.	Incorporación a Didascalia . . . . .	50
4.1.3.	Diseño inicial . . . . .	50
4.1.4.	Transcripción de partes disciplinarios . . . . .	51
4.2.	Desarrollo de la herramienta . . . . .	52
4.2.1.	Reestructuración de la aplicación . . . . .	52
4.2.1.1.	Estructura del proyecto . . . . .	52

4.2.1.2.	Estructura y generación de las situaciones . . . . .	52
4.2.1.3.	Escena Menú . . . . .	54
4.2.1.4.	Escena Class_GameScene . . . . .	54
4.2.1.5.	Escena CapturePose . . . . .	58
4.2.2.	Implementación de la aplicación en Realidad Virtual . . . . .	58
4.2.2.1.	Lógica y animaciones de los estudiantes . . . . .	60
4.2.3.	Incorporación de Emopose . . . . .	61
4.2.3.1.	Problemas durante la incorporación . . . . .	62
4.2.4.	Adición de nuevas emociones . . . . .	63
4.2.4.1.	Investigación sobre el lenguaje corporal y las emociones . .	63
4.2.4.2.	Preparación del equipo . . . . .	67
4.2.4.3.	Creación del Data Set de emociones . . . . .	68
<b>5</b>	<b>Experimentación</b>	<b>70</b>
5.1.	Preparación . . . . .	70
5.2.	Participantes . . . . .	71
5.3.	Diseño experimental . . . . .	71
5.3.1.	Explicación de la prueba . . . . .	71
5.3.2.	Desarrollo de la prueba . . . . .	72
5.3.3.	Cuestionario post-prueba . . . . .	73
5.3.4.	Entrevista final . . . . .	74
5.4.	Correcciones . . . . .	74
<b>6</b>	<b>Resultados</b>	<b>76</b>
6.1.	Resultados recogidos por la aplicación CVR-MC . . . . .	76
6.2.	Resultados cuantitativos de la prueba de usabilidad del sistema . . . . .	78
6.3.	Resultados cualitativos de la entrevista . . . . .	80
<b>7</b>	<b>Discusión</b>	<b>82</b>
7.1.	Limitaciones del estudio . . . . .	83
<b>8</b>	<b>Aportaciones individuales</b>	<b>84</b>
8.1.	Sandra Alonso Paz . . . . .	84
8.2.	Daniel López Acero . . . . .	85
8.3.	Andrés Puente Rodríguez . . . . .	86
8.4.	Antonio Luis Suárez Solís . . . . .	87
<b>9</b>	<b>Conclusiones</b>	<b>88</b>
9.1.	Objetivos cumplidos . . . . .	89
<b>10</b>	<b>Conclusions</b>	<b>91</b>
10.1.	Accomplishment objectives . . . . .	92
<b>11</b>	<b>Trabajo futuro</b>	<b>94</b>
<b>12</b>	<b>Anexo</b>	<b>96</b>
12.1.	Publicación de un paper en el ICCE 2021 . . . . .	96
12.2.	Consentimiento firmado . . . . .	107
12.3.	Preguntas del cuestionario post-test . . . . .	110



# Índice de figuras

2.1. Distribución representativa de la comunicación y sus componentes según Meh- rabian. Imágen de <a href="#">Belén Ávila</a> . . . . .	16
2.2. Ejemplo de <i>wrapper</i> de Kinect. . . . .	21
2.3. Entorno software de Axis neuron. . . . .	22
2.4. Dispositivo Smartsuit Pro colocado sobre un usuario. . . . .	22
2.5. A la izquierda InSight que recoger las fuerzas. A a la derecha Motive 2.2, que recoge movimiento. . . . .	23
2.6. Ejemplo de resultado obtenido por Xnet. . . . .	23
2.7. Dispositivo Neuroelectrics Enobio 32. . . . .	24
2.8. Resultado de ejecución de ML Kit. . . . .	25
2.9. Representación de las 2 redes usadas para el preprocesamiento de imágenes. . . . .	26
2.10. Interfaz de la aplicación EmoPose. . . . .	27
2.11. Estimación emocional realizada por iMotions. . . . .	28
2.12. Dispositivo <i>Eye Tribe Tracker</i> . . . . .	28
2.13. Entorno virtual de Classroom VR. . . . .	30
2.14. Visor y controladores Touch de Oculus Rift S . . . . .	31
2.15. Escenas de la clase virtual para el diagnóstico de Déficit de Atención. . . . .	32
2.16. Reproducción de la escena del crimen mediante Forensic Holodeck. . . . .	33
3.1. Planificación de los meses septiembre, octubre y noviembre . . . . .	39
3.2. Planificación de los meses diciembre y enero . . . . .	39
3.3. Planificación de los meses febrero y marzo . . . . .	39
3.4. Planificación de los meses marzo y abril . . . . .	40
3.5. Planificación de los meses abril y mayo . . . . .	40
3.6. Estructura de la Unidad Compartida Drive. . . . .	41
3.7. Escritorio de Unity . . . . .	42
3.8. Importación del SDK para Perception Neuron en Unity . . . . .	43
3.9. Configuración del setup de streaming de datos en Axis Neuron . . . . .	43
3.10. Ejemplo de ejecución de la herramienta ThreeDPoseUnityBarracuda . . . . .	45
3.11. Componentes de Oculus Rift S . . . . .	46
3.12. Componentes del dispositivo Perception Neuron. . . . .	46
3.13. Pruebas para la lectura del lenguaje corporal realizadas con Kinect . . . . .	47
3.14. Gafas EmteqPRO . . . . .	48
4.1. Primera reunión con el equipo Didascalia (17/12/2020). . . . .	50
4.2. Estructura de Managers de CVR-MC. . . . .	54
4.3. Captura de nuevas posiciones mediante la interfaz de EmoPose. . . . .	58
4.4. Importación al proyecto . . . . .	59

4.5.	Estructura del proyecto . . . . .	59
4.6.	Despliegue de UiHelppers . . . . .	60
4.7.	Esquema en Unity de las animaciones de los alumnos . . . . .	61
4.8.	Funcionamiento de <i>Inverse Kinematics</i> . . . . .	62
4.9.	Unidades de acción: AU1, AU2, AU5, AU26. Emoción expresada: sorpresa. . . . .	65
4.10.	Unidades de acción: AU4, AU9, AU10. Emoción expresada: disgusto . . . . .	65
4.11.	Combinaciones de expresiones corporales asociables una emoción. Alegría, disgusto, ira, miedo, sorpresa, tristeza y neutral respectivamente . . .	66
4.12.	Calibrado del traje <a href="#">Perception Neuron</a> mediante <a href="#">Axis Neuron</a> . . . . .	67
4.13.	Emociones asociadas a la expresión corporal [Schindler et al., 2008] . . . . .	68
4.14.	Esquema final de las emociones que conforman el Data Set de Classroom VR- Motion Capture . . . . .	69
5.1.	Sala Polivalente de la Facultad de Educación de la Universidad Autónoma de Barcelona . . . . .	70
5.2.	Gráficos extraídos del <a href="#">cuestionario post-experimento</a> sobre datos demográficos de los participantes . . . . .	71
5.3.	Esquema de las etapas del experimento . . . . .	74
6.1.	Primera parte del CSV resultado de la simulación del participante 2 (P02) . . .	77
6.2.	Segunda parte del CSV resultado de la simulación del participante 2 (P02) . .	78

# Capítulo 1

## Introducción

### 1.1. Motivación

Contar con docentes cualificados y experimentados al frente de las aulas es imprescindible para el desarrollo de las futuras generaciones en cualquier sociedad. Sin embargo, la realidad es que no todos los docentes han podido disfrutar de una preparación práctica que les permita gestionar situaciones con las que no están familiarizados como, por ejemplo, las situaciones conflictivas.

El principal motivo de la falta de experiencia práctica en este sector, es la ausencia de entornos seguros que permitan poner a prueba estas habilidades de gestión del clima del aula.

A la hora de gestionar situaciones disruptivas, como los conflictos escolares, los docentes deben saber controlar su forma de expresión, es decir, su comunicación verbal y no verbal. Aunque parece evidente que las palabras muestran lo que el cerebro piensa, en la comunicación no verbal no siempre es así. Esto se debe al carácter involuntario de este tipo de comunicación.

La falta de práctica en situaciones realistas, sumado a la escasez de aplicaciones que ayuden a mejorar la comunicación efectiva o asertiva hace que la comunicación no verbal sea una asignatura pendiente para la mayoría de docentes que comienzan su carrera profesional. Sin embargo, es algo que puede enseñarse y que mejora en gran medida las habilidades de resolución de conflictos.

Nuestro objetivo con este trabajo es dar solución a ambos problema utilizando nuevas tecnologías. Por este motivo desarrollamos **ClassRoom VR-Motion Capture** (CVR-MC), una herramienta basada en realidad virtual, que permite a los docentes en formación practicar y mejorar sus capacidades comunicativas en un entorno realista y seguro. CVR-MC también proporciona retroalimentación sobre la toma de decisiones y el lenguaje no verbal del docente durante la simulación. Esto permite, no solo una forma de aprender más segura y efectiva, sino también la posibilidad de crear debates enriquecedores entre los alumnos de docencia y los maestros, sobre qué comportamientos se podría haber llevado a cabo, qué actitudes y estrategias se han empleado y cómo se pueden mejorar.

## 1.2. Objetivos

### 1.2.1. Objetivos generales

El objetivo principal del proyecto es desarrollar y evaluar un entorno de realidad virtual para que los docentes en formación puedan mejorar sus competencias comunicativas de gestión del clima en el aula, en un entorno seguro. A lo largo de la simulación el usuario podrá interactuar con el entorno tratando de gestionar diferentes situaciones que pueden surgir de manera inesperada y que obstaculizan el desarrollo de la clase.

El sistema contará con reconocimiento de emociones, gracias al análisis del lenguaje no verbal, incluyendo el movimiento corporal, palabras clave y entonación empleada. Además llevaremos a cabo una evaluación para verificar el correcto funcionamiento y utilidad del sistema para la formación de profesores noveles.

### 1.2.2. Objetivos específicos

- Implementar un sistema de escenas y situaciones generalizado que permita introducir de forma sencilla nuevos escenarios manteniendo las funcionalidades ya existentes como el reconocimiento de palabras y la extracción de emociones gracias al análisis de la entonación.
- Integrar y evaluar en el proyecto un sistema de reconocimiento de gestos y posturas para clasificar en tiempo real el lenguaje corporal y extraer la emoción predominante del usuario en cada momento. Para ello utilizaremos el traje Perception Neuron y la herramienta EmoPose.
- Integrar el uso de Realidad Virtual en la simulación para permitir mayor inmersión y realismo en las situaciones aumentando así su utilidad para el profesorado.
- Serializar adecuadamente los principales datos obtenidos (posicionamiento, captura de movimiento corporal, vocabulario utilizado y decisiones tomadas durante la situación) para poder analizarlos y sacar conclusiones de ellos.
- Evaluar de manera adecuada el feedback aportado para ser de la máxima utilidad al profesorado en experiencias con alumnos reales.
- Realizar una prueba piloto en colaboración con la Universidad Autónoma de Barcelona y más tarde, analizar los resultados obtenidos y extraer conclusiones de los mismos. Más concretamente sobre su usabilidad y efectividad en la formación de docentes noveles.
- Con el fin de hacer pública nuestra investigación y el desarrollo de la herramienta, escribiremos un artículo y lo presentaremos al congreso *29th International Conference on Computers in Education. Asia - Pacific Society for Computers in Education (ICCE 2021)*



### 1.3. Estructura del documento

Este documento se divide en doce capítulos, cada uno enfocado en una temática distinta. Este epígrafe está enmarcado en primer capítulo del documento llamado [Introducción](#) donde además se explica la motivación y los principales objetivos de este trabajo.

En capítulo [Estado del arte](#) se muestran los resultados de los estudios e investigaciones iniciales de los antecedentes de este proyecto. Entre sus apartados, describimos los conceptos que sientan la base de este trabajo. Estos son: la comunicación no verbal y su importancia en la transmisión de conocimiento y la realidad virtual y sus ventajas e inconvenientes a la hora de utilizarla en procesos de formación. Además, se detallan las técnicas actuales que se utilizan en el manejo de conflictos en el aula y cómo los docentes aprenden hoy en día este tipo de metodologías. Por último, hablamos de las tecnologías aplicadas a la captura y análisis del lenguaje no verbal y algunos de los proyectos en los que se han utilizado.

En el capítulo [Metodología y tecnología](#) describimos primeramente, la metodología empleada para el desarrollo del proyecto, las funciones de cada miembro del equipo y la planificación general. A continuación explicamos las diferentes tecnologías, software y dispositivos utilizados y propuestos a la hora de desarrollar la herramienta.

Seguidamente, en el capítulo [Diseño y desarrollo de la herramienta](#) se describe el proceso de diseño e implementación del proyecto. Este comprende desde la decisión de la idea principal del Trabajo de Fin de Grado (fase divergente) y diseño inicial, hasta la creación de un Data Set de emociones.

El quinto capítulo, [Experimentación](#), describe las pruebas finales llevadas a cabo en la Universidad Autónoma de Barcelona, incluyendo el diseño experimental y las correspondientes correcciones de la herramienta para su mejora.

En el capítulo [Resultados](#), se muestran los resultados cuantitativos y cualitativos obtenidos de la experimentación citada anteriormente. A continuación, en el capítulo [discusión](#) se hace una reflexión sobre los mismos.

A lo largo del capítulo [Aportaciones Individuales](#) se despliega el valor aportado por cada uno de los componentes del equipo.

A continuación se obtienen las conclusiones finales del trabajo y se reflejan en el capítulo [Conclusiones](#). A este capítulo, le sigue [Trabajo futuro](#), donde reflejamos los siguientes pasos a seguir para la mejora de esta herramienta.

Por último, se incluyen el [Anexo](#) y las [Referencias Bibliográficas](#) en los dos últimos capítulos del documento.

## Capítulo 2

# Estado del arte

### 2.1. La comunicación

La comunicación humana es considerada la manera de construir y compartir conocimiento y de crear y agrupar comunidades y culturas. Este concepto fundamental ha sido y es estudiado por reconocidos profesionales de distintas disciplinas como el filósofo Karl Jaspers [Rodríguez, 1984] que dedicó su carrera al estudio de los problemas principales del ser humano como la comunicación, el lingüista estadounidense Benjamin Lee Whorf quien trató de averiguar cómo la lengua y la comunicación modelan el pensamiento [Parra, 1988], o el sociólogo Max Horkheimer que junto a Theodor Adorno desarrolló *Dialéctica de la Ilustración* (1947).

Según el diccionario de la [Real Academia Española], la comunicación es la transmisión de señales mediante un código común al emisor y al receptor. Estos son los principales componentes de la comunicación, sin embargo, existen más.

La comunicación está compuesta por distintos elementos [Uribe 2009], los principales son: el emisor o persona que emite el mensaje, el receptor o persona que recibe el mensaje, el mensaje o contenido de la información que se envía, el canal o medio por el que se transmite el mensaje, el código o signos y reglas empleadas para enviar el mensaje y por último el contexto o situación en la que se produce la comunicación.

Las formas de comunicación humana pueden agruparse en dos grandes categorías, la comunicación verbal y la no verbal.

La comunicación verbal es aquella en la que se utilizan las palabras [Juncosa, 1999] y el mensaje se transmite y expresa a través de la comunicación oral o escrita. Por otra parte, la comunicación no verbal se refiere a la acción de comunicar sin hablar y se asocia a la inteligencia emocional [Cestero Mancera, 2014].

En el presente trabajo exploraremos el campo de la comunicación no verbal y sus principales componentes. Más concretamente, analizaremos la expresión no verbal de los docentes en formación con el fin de mejorar sus habilidades comunicativas en el aula. A continuación se detalla la definición de comunicación no verbal y su importancia a la hora de transmitir mensajes.

### 2.1.1. Comunicación no verbal

Para llegar a entender, de manera global, el alcance que tiene la comunicación no verbal en los procesos comunicativos es necesario hacer hincapié en los componentes esenciales de la comunicación no verbal [Vargas, 2016], estos son la proxémica, la paralingüística y la kinésica. De este modo, podremos demostrar el papel fundamental que tienen en la transmisión de mensajes.

La comunicación no verbal paralingüística se compone de signos orales, auditivos, táctiles y visuales. Los elementos paralingüísticos son las expresiones de sonidos sin palabras como por ejemplo la entonación, gruñidos, bostezos, el ritmo de comunicación, etc. Autores como [Poyatos, 1970] consideran, dentro de la paralingüística, la entonación como el elemento más distintivo. Llegando clasificar sus cualidad primarias en las siguientes: timbre, tono, dirección tonal, intervalo silábico, volumen, duración silábica, ritmo y tiempo empleado en terminar la comunicación.

Por otra parte, la componente proxémica se refiere a las distancias entre el emisor y el receptor del mensaje. Partiendo de la definición que acuñó el antropólogo Edward T. Hall en 1963, varios autores como [García, 2000] y [Schmidt, 2013], afirman la existencia de cuatro tipos de distancias interpersonales: la distancia íntima (0-60 cm), la personal (6-120 cm), la social (120-300 cm) y la pública (más de 300 cm).

Por último, la componente kinésica de la comunicación no verbal o lenguaje corporal corresponde a los gestos corporales y las miradas. Esta es la componente que más varía según el lugar (continente, país, ciudad, etc) donde se interprete debido a los claros rasgos culturales que la sostienen [Kendon, 1972]. Algunos ejemplos de este tipo de comunicación son: levantar las cejas, posición de los brazos, la postura o la posición con respecto al receptor del mensaje.

#### 2.1.1.1. La importancia de la comunicación no verbal

Como adelantamos en el apartado anterior, la comunicación, tienen una gran componente no verbal. Varios autores como los antropólogos Edward T. Hall (1959) [Hall, Edward. T. , 1959] y Ray Birdwhistell [Ponce, 1980], realizaron estudios sobre la importancia del lenguaje no verbal en la comunicación. Sin embargo, el estudio más relevante en este campo es el de Albert Mehrabian, un reconocido psicólogo e investigador de los años 80 que dedicó gran parte de su carrera al estudio de actitudes y sentimientos relacionados con la comunicación no verbal. Estos estudios le llevaron a desarrollar varios experimentos entre los que destaca el de la enunciación de la Regla de Mehrabian [Mehrabian, 1971].

La regla de Mehrabian o más conocida como la regla 7-38-55 hace referencia a los pesos asignados a cada uno de los componentes de la comunicación durante la transmisión de un mensaje oral [Pereiro, 2019]. El primero de ellos hace referencia a la componente verbal, que contribuye un 7 % en esta tarea. Estas son básicamente las palabras y conectores que utilizamos para transmitir el mensaje. Por otra parte, el 38 % corresponde a la componente paralingüística que, como hemos explicado anteriormente, comprende todos los signos orales, auditivos, táctiles y visuales. Dado que estamos transmitiendo un mensaje oral, gran parte de este porcentaje corresponde a la voz, su volumen, entonación y tonalidad. Por último encontramos la componente kinésica, el 55 %. Esta es la que mayor peso tiene a la

hora de transmitir un mensaje oral y es la referida al lenguaje corporal, es decir, gestos mirada, postura, etc.

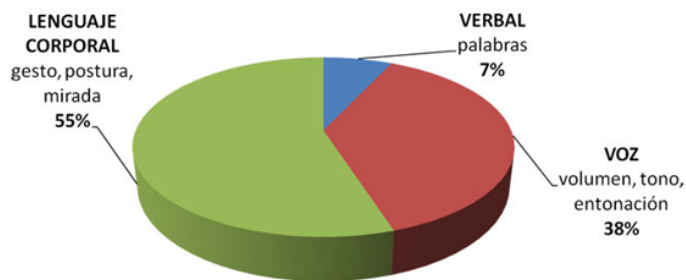


Figura 2.1: Distribución representativa de la comunicación y sus componentes según Mehrabian. Imágen de [Belén Ávila](#)

Autores posteriores como el psicólogo Carlos J. Hofstadt Román [Van der Hofstadt Román, 2005], [Orzáiz, O. , 2009] o la doctora en Filología Hispánica Susana Ridaó [Ridaó Rodrigo, 2017] han confirmado mediante varios experimentos que esta regla se cumple ampliamente. Sin embargo, otros autores como [Yaffe, 2011] son más precavidos y aunque confirman la veracidad de la regla, concluyen que sólo puede ser utilizada en determinados estudios, más concretamente los dedicados al análisis de actitudes o sentimientos mediante la comunicación no verbal.

#### 2.1.1.2. Comunicación no verbal y las emociones

Una vez introducida la comunicación no verbal, sus componentes y la importancia de los mismos en la transmisión de mensajes entre emisor y receptor, concluimos que si bien los estudios de Mehrabian son ciertos, estos son utilizados principalmente para la extracción de actitudes y emociones a través del lenguaje no verbal.

El psicólogo, antropólogo y sociólogo Paul Ekman tomó como base los estudios de Mehrabian y, basándose en que el 55 % de la comunicación está representado por el lenguaje no verbal, realizó varios estudios en este campo.

A pesar de que antropólogos anteriores a él pensaban que la expresiones corporales y faciales están determinadas principalmente por la cultura en la que se desarrolla el individuo, Ekman estableció que estas expresiones son universales [Ekman & Davidson, 1994] y que tienen un origen biológico cómo explica Darwin en la Teoría de la evolución de las especies [Darwin et al., 1971].

Ekman realizó varios experimentos con tribus aisladas del Estado Independiente de Papúa Nueva Guinea. Les mostró varias imágenes recogidas de individuos de distintas culturas expresando facial y corporalmente una emoción. Los miembros de esta tribu fueron capaces de reconocer las emociones que pretendían transmitir. Este experimento concluyó que, aunque esas personas no habían tenido contacto con otras culturas, y por lo tanto no estaban influenciados, eran capaces de descifrar la comunicación no verbal de otros individuos.

A raíz de este estudio, Ekman describió una lista de emociones básicas (1972) que, aún a día de hoy, sirven como referencia para multitud de estudios actuales no solo relacionados con la psicología y sociología [Levav, 2005] sino también con carreras técnicas como informática [Normoyle et al., 2013]. Estas emociones son: alegría, ira, miedo, asco, sorpresa y tristeza. Aunque más tarde en 1990, Ekman añadió más emociones a esta lista como alivio, soberbia o vergüenza, nos son consideradas como básicas.

## **2.2. Docencia**

Los docentes cumplen un papel fundamental en la sociedad debido a que su trabajo va más allá de enseñar materias; su papel de transformación social, de transmisión de valores y de formación de personas críticas y autónomas es fundamental. Por eso es indispensable que cuenten con una buena formación, que deberá ir actualizándose a lo largo de toda su carrera profesional.

En España, como en la mayoría de países de la Unión Europea, la enseñanza está dividida en cuatro etapas: educación infantil (0-6 años), primaria (6-12 años), secundaria (12-16 años) y superior (16 o más). [Prats et al., 2005]

El sistema educativo actual regula la formación que deben tener los docentes para poder impartir clases en cada una de las etapas escolares mencionadas anteriormente.

### **2.2.1. Educación infantil y primaria**

Los maestros de educación infantil y primaria, encargados de la docencia de niños de cero a seis años y de seis a doce años respectivamente, deben obtener un Grado de Educación Infantil o Primaria, o contar con un título anterior equivalente a efectos de docencia [Beraza & Cerdeiría, 2011] [Rebolledo Gámez, 2015], que les habilita para trabajar en centros públicos, privados y concertados.

El contenido principal de ambos grados es formarse en el desarrollo biológico y psicológico de los alumnos en la etapa correspondiente y sus procesos de aprendizaje. Lo que les confiere amplios conocimientos sobre el tipo de alumnado que tendrán en sus aulas y cómo deben forjar relaciones con ellos. Además, adquieren metodologías de enseñanza y cómo aplicarlas en un entorno donde hay diferencias de aprendizaje entre sus oyentes.

### **2.2.2. Educación secundaria**

En el caso de los profesores de las siguientes etapas, es diferente. Para ser docente de Educación Secundaria se necesita tener un título de grado o licenciatura. Este puede ser Biología, Lengua, Historia del Arte, Matemáticas, etc. En definitiva, de cualquier asignatura que se imparte durante la Educación Secundaria. Además, el licenciado debe contar con un “Máster Oficial en Formación del Profesorado de ESO, Bachillerato, Formación Profesional y Enseñanza de Idiomas” [Muñoz, 2009].

En este caso nos encontramos ante un perfil profesional con los conocimientos científicos necesarios para dominar por completo la materia sobre la que va a impartir su clase. Pero sin apenas experiencia o conocimientos del marco de la docencia. Según un artículo del

ABC [ABC] esto sucede por el escaso contenido que se puede impartir durante el Máster universitario habilitante.

Por este motivo, los profesores principiantes normalmente no se han enfrentado nunca con un grupo de alumnos a los que transmitir sus conocimientos. Nunca ha tenido que explicarse de forma que personas no expertas en el tema que debe exponer lo entiendan, lo asimilen y lo aprendan. De la misma forma, no han tenido que mantener la disciplina, resolver conflictos entre alumnos o favorecer un entorno adecuado para el aprendizaje.

Esta inexperiencia favorece la aparición de cierto sentimiento de inseguridad al encontrarse frente al alumnado. Si el alumno percibe esta inseguridad, y la interpreta como falta de conocimiento es muy probable que se fomente su desmotivación y pérdida de la figura de autoridad que es el profesor en el aula. Este tipo de situaciones propician el surgimiento de conflictos.

En general, se crea un círculo vicioso de incomprensión mutua ya que el docente percibe la actitud reacia de sus alumnos y entra en un estado de mayor inseguridad, haciéndole incapaz de resolver ningún tipo de situación que se de en la clase.

## **2.3. Gestión del aula**

Tras la consulta de varias fuentes como [Bentinck, 2018] o [Emmer & Stough, 2001], la gestión del aula puede definirse como las acciones de los docentes para mantener la atención de los estudiantes, y a través de ello, crear un entorno que facilite el logro de los objetivos instructivos y el bienestar socioemocional del alumnado. De acuerdo con [Doyle, 2013], la clave del éxito de un maestro en la gestión es, en primer lugar, su conocimiento sobre los eventos que pueden ocurrir en el aula y, en segundo lugar, su habilidad para monitorizar y guiar actividades en base a esta información.

Sin embargo, a pesar de los esfuerzos de los maestros para crear y mantener un ambiente de aprendizaje productivo durante la clase, en el aula ocurren eventos inesperados que interfieren las actividades instructivas. Por tanto, como afirma [Doyle, 2013], el éxito de la gestión del aula depende esencialmente de la competencia de los docentes para comprender e interpretar eventos conflictivos, en circunstancias inmediatas.

### **2.3.1. Gestión de conflictos**

#### **2.3.1.1. Qué es un conflicto**

Antes de describir el estado actual de los conflictos a nivel escolar en nuestro país, es preciso definir y matizar el concepto de conflicto. Según [Vallejo, 2005], “un conflicto se produce cuando hay un enfrentamiento de los intereses o las necesidades de una persona con los de otra, o con los del grupo, o con los de quien detenta la autoridad legítima”.

Este tipo de situaciones son más propensas a darse en entornos cerrados y con roles diferenciados en cuanto a madurez, edad o responsabilidades. Un ejemplo claro que recoge todas estas características es el ambiente escolar.

Varios autores como Pérez y Pérez de Guzmán [Pérez-de-Guzmán et al., 2011] definen

conflicto en el aula como “el desacuerdo existente entre personas o grupos en cuanto a ideas, intereses, principios y valores, dentro de una comunidad escolar, de tal manera que las partes perciben sus intereses como excluyentes, aunque puedan no serlo.”

#### **2.3.1.2. Formación del profesorado en gestión de conflictos**

A pesar de la indudable importancia del tema, de acuerdo con informes recientes, existe un reclamo urgente para cubrir carencias formativas de los docentes españoles respecto a la competencia de gestión del clima de aula. Según [Schleicher, 2020], España se encuentra muy por debajo de la media de profesores que refieren haber recibido formación sobre la gestión de los comportamientos conflictivos de los estudiantes y el clima de aula. Tan solo un 40 % afirmaron sentirse preparados para controlar una clase.

Asimismo, un estudio reciente sobre la satisfacción de la formación recibida por estudiantes en el Máster de formación para el profesorado realizado en cuatro importantes universidades españolas, concluyó que, al iniciar del máster, los estudiantes afirmaron no tener apenas conocimiento de los procesos de interacción y comunicación en el aula y el dominio de habilidades necesarias para fomentar el aprendizaje, la convivencia y abordar problemas de disciplina y resolución de conflictos de aula. Al finalizar el máster, solo el 34 % se situaban en la valoración más alta, frente al 19,8 % alcanzada inicialmente [Sarceda-Gorgoso et al., 2020].

#### **2.3.2. Importancia de aprender a gestionar el aula**

El profesor debe asumir el protagonismo como organizador y director de toda actividad en el aula y por lo tanto ser capaz de tener el control de cualquier situación que se pudiese producir en la misma.

El problema principal entre los docentes noveles reside en su inexperiencia en el tratamiento de situaciones en el aula. Particularmente debemos tener en cuenta aquellos que se dedican a la enseñanza en Educación Secundaria y que, como se ha comentado en el apartado [Educación secundaria](#), aunque tienen amplios conocimientos en los conceptos que imparten, no han desarrollado completamente su capacidad a la hora de tratar con alumnos.

Las investigaciones de autores como [Bentinck, 2018] muestran que los profesores que deben enfrentarse a situaciones conflictivas, es porque ha fallado tanto la planificación como la aplicación de sus habilidades para el manejo y control del aula. Es por tanto, vital que los docentes noveles disfruten de un programa formativo orientado a al mantenimiento de un buen clima del aula.

#### **2.3.3. Importancia de la comunicación en el manejo del aula**

Como hemos mencionado en apartados anteriores, la comunicación no verbal está presente en toda interacción humana. Podemos deducir por tanto, que la educación no es una excepción a esta regla.

Según autores como [Gutiérrez Fernández, 2007], el comportamiento no verbal de un docente al frente de una clase puede condicionar el ambiente del aula. Es deber del profesor, ser consciente de los mensajes que está enviando al alumnado no sólo a través de

sus palabras, sino también a través de su lenguaje no verbal y procurar que estos no se contradigan, es decir, procurar que nuestra comunicación sea efectiva.

Debemos tener en cuenta que a través del lenguaje corporal se contagian los estados emocionales [Shablico, 2012] y se logran transmitir en el aula actitudes, emociones, sentimientos, valores, prejuicios, estados de ánimo, etc. Y por tanto es de suma importancia adoptar un lenguaje apropiado y acorde a la situación.

Algunos ejemplos de posturas y gestos estudiados por autores como [Mur, 2008] y [Shablico, 2012] y que se deben fomentar son: inclinarse hacia delante cuando se pretende mostrar interés, adoptar una postura erguida para reafirmar seguridad y confianza en lo que se dice o hacer gestos ilustrativos para contagiar entusiasmo.

Además, en lo relativo a imponer autoridad se debe intentar focalizar la mirada allá donde se requiera. A la hora de dar órdenes al alumnado se recomienda ilustrar con gestos firmes y secos lo que se debe hacer. En determinadas situaciones, normalmente donde empieza a predominar el descontrol en el aula, es bueno quedarse en silencio, golpear ligeramente la mesa o incluso dar un pequeño toque en el brazo de la persona despistada. Esto hará que se centre toda la atención de los oyentes en el docente.

Por el contrario, se deben evitar las posiciones tensas o rígidas que evidencian temor o que la persona ha adoptado una posición defensiva y que por lo tanto hacen que el alumno se sienta poderoso y con el control de la clase.

## **2.4. Tecnologías y comunicación no verbal**

Las nuevas tecnologías son utilizadas en todas las ramas del conocimiento y el estudio de la comunicación no verbal no es una excepción. Como se ha mencionado anteriormente en el apartado [La importancia del lenguaje no verbal](#), según el profesor Albert Mehrabian, el 55 % de la comunicación está determinado por el lenguaje corporal (gestos, posturas, respiración, etc.), el 38 % por el tono de voz (entonación, proyección, etc.) y solo el 7 % por las palabras que pronunciamos [Romo, n.d.].

A lo largo de los años se han desarrollado diversos tipos de tecnologías para medir, analizar e interpretar esta información que a menudo pasa desapercibida a la vista humana. Algunas de estas herramientas tecnológicas nos permiten interpretar, en cierta medida, el estado emocional del usuario según su expresión corporal.

### **2.4.1. Herramientas que permiten la captura del lenguaje corporal**

#### **2.4.1.1. Kinect**

Se trata de un controlador creado por Alex Kipman y desarrollado por Microsoft para la consola Xbox 360. Este dispositivo permite a los usuarios controlar e interactuar con la consola sin tener contacto físico con un mando clásico. Esto lo hace mediante una interfaz de usuario que reconoce gestos, voz, objetos e imágenes. Internamente, Kinect detecta los puntos clave como hombros, cabeza, manos, codos, rodillas y pies por lo que ha sido utilizado en diversos estudios en los que era necesario capturar el movimiento corporal,



como *Recognition of Emotions using Kinects* [Li et al., 2015] llevado a cabo por Shun Li y Changye Zhu, de la *Chinese Academy of Sciences*.

En 2010 el español Héctor Martín desarrolló mediante métodos de ingeniería inversa un controlador para GNU/Linux mediante el cual hacía uso de las características de Kinect con una cámara RGB. Si bien el proyecto de código abierto no es accesible a día de hoy de manera sencilla, podemos ver su uso y acceder a él mediante proyectos como el desarrollado por el propio Héctor Martín, que implementó un sistema de alarma doméstica basado en Kinect. Este trabajo puede ser consultado en su [repositorio]

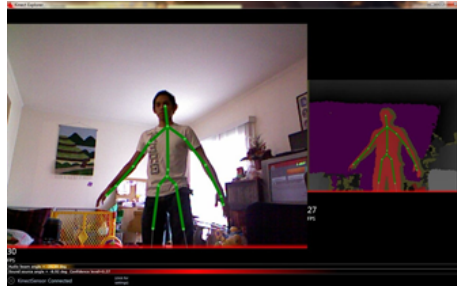


Figura 2.2: Ejemplo de *wrapper* de Kinect.

#### 2.4.1.2. Axis Neuron

Axis Neuron es la aplicación software mediante la cual se conecta el traje de [Perception Neuron](#) al ordenador. Esta muestra una plantilla de las neuronas disponibles en el traje en la forma de un maniquí, indicando por código de colores el estado de la neurona.

La aplicación recibe y procesa la información del movimiento para que pueda ser exportado a software de terceros, como Blender, [Unity](#), Maya u otros, aportando el movimiento propio de una producción de alta calidad.

Su eficacia y fiabilidad han quedado demostradas en diversos estudios, como el publicado por *Journal of Clinical Medicine* por Hyun Soo Kim [Kim et al., 2019], en el que este software junto con el traje Perception Neuron fueron utilizados con el objetivo de corregir los fallos posturales de cirujanos en formación durante sus prácticas para ayudar a mejorar la economía del movimiento y el tiempo con resultados favorables.

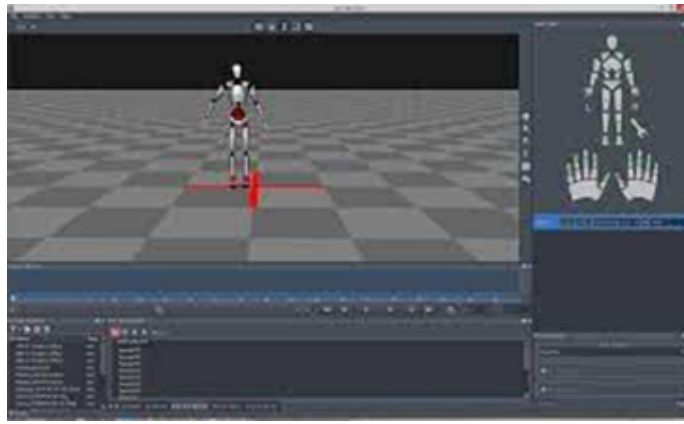


Figura 2.3: Entorno software de Axis neuron.

#### 2.4.1.3. Smartsuit Pro

Se trata de un traje de captura de movimiento desarrollado y distribuido por [Rokoko Electronics](#). El traje incorpora 19 sensores de movimiento distribuidos por todo el cuerpo en puntos clave para capturar el movimiento, como la cabeza, cintura, codos, manos y pies.

Este dispositivo soporta formatos *FBX BVH* y *CSV* y está diseñado para poder enviar a altas velocidades los datos sobre el movimiento del usuario en tiempo real a aplicaciones como Rokoko Studio. Esta aplicación es la encargada de capturar el movimiento, contando con numerosos *plugins* para programas como Blender, Unreal Engine o [Unity](#).



Figura 2.4: Dispositivo Smartsuit Pro colocado sobre un usuario.

#### 2.4.1.4. OptiTrack

[OptiTrack](#) es uno de los mayores proveedores de hardware y software de captura de movimiento en todo el mundo. Entre sus clientes encontramos grandes productoras que han querido incluir sus tecnologías en algunas de las nuevas y espectaculares producciones de Hollywood. Algunos de sus productos son Motive 2.2 y aplicaciones de análisis del movimiento como InSight, que recibe los datos del traje mencionado transmitiendo la posición, la rotación e incluso las fuerzas que están siendo ejercidas.

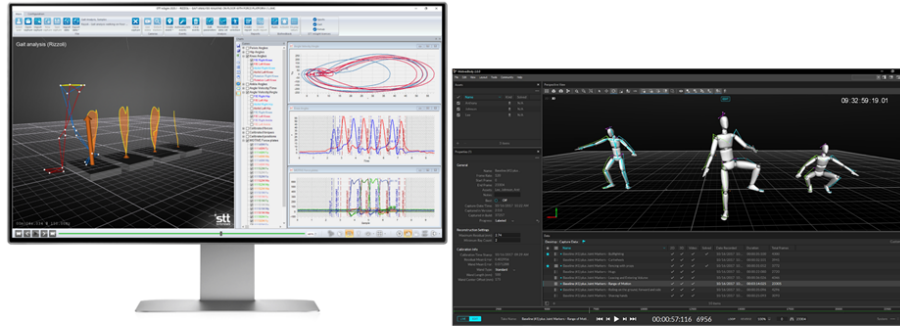


Figura 2.5: A la izquierda InSight que recoge las fuerzas. A a la derecha Motive 2.2, que recoge movimiento.

#### 2.4.1.5. Xnet

Este proyecto, en formato de librería de C++, supone un interesante acercamiento a la captura de movimiento 3D de varias personas simultáneamente y fue presentado en la convención [SIGGRAPH 2020](#).

Su funcionamiento se basa en distintas etapas: En una primera fase, la red neuronal convolucional que posee, identifica las partes individualmente, los huesos. Una vez identificado lo señala como parte móvil y lo asigna a una parte del esqueleto. A continuación, la red neuronal clasifica las distintas partes según si corresponden al mismo individuo o no, diferenciando así cuantas personas hay en escena y aproximadamente como están dispuestas. De esta manera crea una pose en 2D y otra en 3D de cada individuo. Para finalizar, se aplican ajustes en la relación espacio-temporal a las predicciones 2D y 3D de cada sujeto, para lograr la mayor coherencia posible y devolver un esqueleto 3D completo con la auténtica pose del usuario.

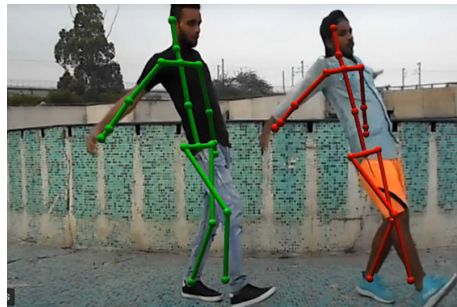


Figura 2.6: Ejemplo de resultado obtenido por Xnet.

#### 2.4.1.6. Neuroelectrics Enobio 32

[Neuroelectrics Enobio 32](#) se trata de un casco inalámbrico con sensores electrofisiológicos repartidos por su superficie, que permiten realizar y almacenar los datos de electroencefalogramas, mientras envía los datos a un ordenador conectado por Wi-fi o Bluetooth, con un tiempo de uso continuado de 14h.

La utilización de este tipo de dispositivos puede ser muy relevante, como se puede observar en estudios como *Assessing the emotional impact of virtual reality-based teacher training* [Stavroulia et al., 2019] en el que midieron mediante dispositivos similares, la presencia de ondas alfa (asociadas a un estado de relajación) y ondas beta (asociadas con un estado de estrés cognitivo) en distintos tipos de situaciones. Con estos datos pudieron observar el grado de estrés al que estaba siendo sometido el usuario y si existía algún desencadenante generalizado. Este modelo en particular ha sido usado en estudios mas cercanos a la medicina, como *Computational Intelligence Techniques for Electrophysiological Data Analysis* [Riera Sardà, 2012], desarrollado por Alejandro Riera de la Universidad de Barcelona , cuyo objetivo era demostrar la eficacia de distintas técnicas computacionales que permiten el correcto análisis de electroencefalogramas y pruebas similares.



Figura 2.7: Dispositivo Neuroelectrics Enobio 32.

#### 2.4.1.7. ML Kit

Se trata de una SDK (software development kit) ofrecida por Google que ofrece acceso a un amplio abanico de modelos de aprendizaje automático, entre otros el reconocimiento de textos, detección de rostros, respuestas inteligentes y más, que pueden ser usados en dispositivos Android.

En referencia al campo que nos ocupa en este apartado, la detección de rostros, puede llevarse a cabo de dos formas, de manera offline en tu dispositivo Android o en la nube, que dependiendo de la frecuencia del uso puede ser más costoso. El modelo *Face Detection* es capaz de reconocer y localizar gestos faciales, obtener coordenadas de ojos, cejas, labios y nariz, así como diferenciar cuando una persona sonríe o tiene los ojos cerrados. Entre sus capacidades también está ser capaz de mantener la identificación única de cada uno de los participantes en vídeo a tiempo real a velocidad suficiente como para poder usarse en aplicaciones de manipulación de vídeo.

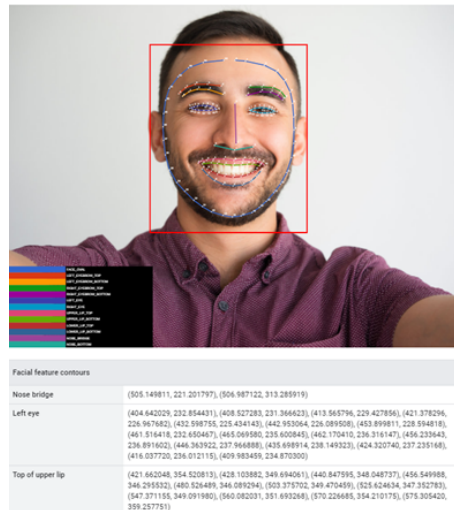


Figura 2.8: Resultado de ejecución de ML Kit.

## 2.4.2. Herramientas que permiten su análisis

### 2.4.2.1. Active Appearance models (AAM)

Un AAM es un algoritmo incluido dentro del campo de la visión artificial, un área de la informática cuyo objetivo es la recepción, procesamiento y análisis de imágenes del mundo real para poder traducirla a información numérica o simbólica interpretable por un ordenador. Su principal característica es la de poder reconocer una imagen y asociarla con un modelo estadístico de su forma previamente generado, con buenos resultados.

Este tipo de modelos se crean en la fase de entrenamiento, mediante una base de datos de imágenes con las coordenadas de los puntos de interés, que se envían al supervisor del entrenamiento. El algoritmo usa las diferencias entre la estimación que facilita el modelo estadístico y la imagen recibida como entrada para optimizar el proceso, seleccionando los puntos de mayor interés, y aprovechando la técnica de los mínimos cuadrados (una técnica de análisis numérico que, dado un conjunto de pares ordenados de la forma variable independiente - variable dependiente y una serie de funciones, trata de encontrar la función continua que aproxime los datos de acuerdo según el criterio de mínimo error cuadrático), lo que permite que este tipo de modelos reconozcan imágenes de manera sencilla y rápida.

Este tipo de aproximación al análisis del lenguaje corporal puede verse en estudios como *Face Recognition Using Active Appearance Models* [Edwards et al., 1998], llevado a cabo por Gareth J. Edwards, Timothy F. Cootes y Christopher J. Taylor, de la Universidad de Manchester, cuyo objetivo es utilizar este tipo de modelos para reconocer los rostros con distintos gestos y expresiones faciales con resultados muy satisfactorios.

### 2.4.2.2. Redes Neuronales Convolucionales

Las redes neuronales convolucionales (CNN, Convolutional Neural Networks) son un tipo de red neuronal artificial que se diferencia de otras por la forma de procesamiento de sus capas, similar a como se procesan en la corteza visual primaria, para detectar e identificar características variadas en la entrada. Las capas se organizan jerárquicamente.

Las primeras detectan formas más simples, como líneas, hasta capas más especializadas que pueden reconocer rostros o formas. Estos tipos de redes neuronales suponen un gran activo en tareas de visión artificial, como el reconocimiento de imágenes.

Su uso para el reconocimiento de imágenes para distintas áreas se ha demostrado en los últimos años en más de un estudio, como en *Sign Language Recognition Using Convolutional Neural Networks*, realizado por Lionel Pigou, Sander Dieleman Pieter-Jan Kindermans y Benjamin Schrauwen, de la Universidad de Ghent. En su proyecto lograron reconocer de manera exitosa 20 signos pertenecientes al [lenguaje de signos italiano](#) aplicando técnicas de preprocesamiento con 2 CNN y usando esta información para relacionarlo con su traducción en una base de datos, con el objetivo de facilitar la comunicación e integración de las personas con discapacidades comunicativas.

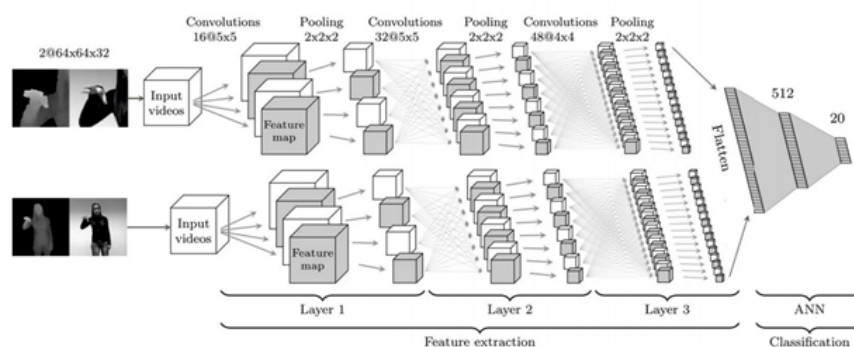


Figura 2.9: Representación de las 2 redes usadas para el preprocesamiento de imágenes.

### 2.4.3. Herramientas que extraen emociones del lenguaje corporal

#### 2.4.3.1. Emopose

Se trata de una aplicación desarrollada por Iván García-Magariño [García-Magariño et al., 2019] del Grupo de Investigación en *Aplicaciones Sociales e Interdisciplinarias basados en Agentes* como autor principal. Su objetivo principal es permitir a los usuarios tomar mayor consciencia de sus emociones y ayudar a los profesionales relacionados a tener en cuenta sus necesidades psicológicas.

La aplicación funciona mediante la manipulación de la postura corporal de un avatar, incluyendo pies, rodillas, caderas, hombros, codos y cuello. La manipulación resulta muy sencilla e intuitiva de realizar. La aplicación permite también el movimiento de las manos para configurar distintas posiciones o aperturas de estas, que están altamente relacionadas con el estado emocional.

Una vez que la postura es la deseada por el usuario, el programa estima la emoción que probablemente esté teniendo según la posición corporal. La estimación se realiza consultando una serie de poses que el programa guarda con una emoción ya asignada, y mediante el método del vecino más cercano calcula cual es la más próxima y por tanto la más probable. El individuo puede confirmar la veracidad de la estimación para aportar retroalimentación

a la aplicación.

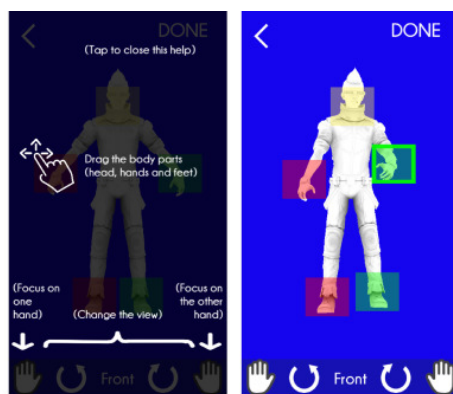


Figura 2.10: Interfaz de la aplicación EmoPose.

#### 2.4.3.2. iMotions

Se trata de un módulo desarrollado por Afectiva Solutions, llamado inicialmente Emotion SDK, y actualmente integrado dentro del proyecto [iMotions](#)

Este módulo puede detectar emociones en tiempo real mediante diversos tipos de sensores ópticos, cámaras estándar o cámaras web, llegando a ser capaz de identificar 7 emociones, 20 expresiones y 13 emoticonos distintos.

Los algoritmos de visión identifican los puntos clave en la cara, como los labios, las cejas, la punta de la nariz, etc. Los algoritmos de clasificación analizan los píxeles de las regiones señaladas para clasificar las expresiones faciales basado en el [Facial Action Coding System](#) en distintas Unidades de Acción.

El Facial Action Coding es un sistema para denominar movimientos faciales humanos por su apariencia en la cara, basado en un sistema desarrollado originalmente por un anatomista sueco llamado Carl Herman Hjortsjö. Se trata de un estándar común de clasificar sistemáticamente la expresión física de las emociones y ha demostrado ser de utilidad para la psicología y la animación. Debido a problemas de tiempo y de percepción subjetiva, la codificación facial se ha establecido como un sistema automático realizado por ordenador que detecta los rostros, características geométricas de las caras y produce perfiles temporales de cada movimiento de la cara.

Puede usarse en conjunto con otras tecnologías como [Kinect](#), como hicieron Juan Sebastián Hernández y Liliam Camila Diosa, del Grupo de Investigación DIGITI, de la Universidad Distrital Francisco José de Caldas, en su trabajo *Sistema de Detección de Emociones a partir de la posición corporal utilizando el sensor Kinect* [Moya Hernández & Diosa Silva, n.d.], en el que se propusieron automatizar la detección de emociones que se asocian a una posición corporal mediante el uso de Kinect y del señalado módulo. El software, en última instancia, está diseñado para poder utilizarse con la gama de productos de iMotions como sensores para medir la conductividad eléctrica de la piel, el pulso, o las corrientes eléctricas que circulan por el sistema nervioso, con la finalidad de otorgar más precisión en la



predicción.

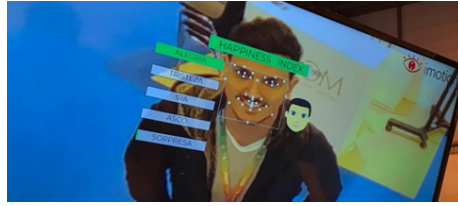


Figura 2.11: Estimación emocional realizada por iMotions.

#### 2.4.4. Eye-Tracking Analysis for Emotions Recognition

Este proyecto desarrollado por Pawel Tarnowski, Marcin Kolodziej, Andrzej Majkowski y Remigiusz Jan Rak [Tarnowski et al., 2020], de la Universidad tecnológica de Varsovia, consiste en un sistema que, mediante la obtención de métricas del ojo, es capaz de detectar la emoción que transmite el usuario.

Las métricas oculares las obtienen por medio del dispositivo *Eye Tribe Tracker*, desarrollado por [The Eye Tribe], una empresa danesa enfocada al desarrollo de dispositivos de seguimiento ocular. Las detecciones son obtenidas gracias al uso de una cámara y un emisor de infrarrojos, que permiten ubicar en todo momento los ojos del usuario, con una distancia efectiva de entre 45 a 75 cm, con una velocidad de respuesta de 20 milisegundos.

Para la clasificación de las métricas y la relación con las emociones emplearon análisis discriminante lineal, máquinas de vectores de soporte (SVM) y un último clasificador basado en el método de “los k vecinos más próximos”. Los parámetros de cada clasificador se eligieron experimentalmente basándose en la eficacia de cada uno, teniendo 25 iteraciones en la fase de entrenamiento.

Su metodología fue mostrar a los participantes, individualmente, una serie de películas que pudiesen despertar emociones en ellos mientras el sistema de seguimiento ocular recoge la información de la posición, la dilatación pupilar y el ritmo de parpadeo, para su posterior análisis y detección de emociones



Figura 2.12: Dispositivo *Eye Tribe Tracker*.



## **2.5. Entornos automáticos de aprendizaje**

### **2.5.1. Videojuegos educativos (serious games)**

#### **2.5.1.1. Qué son**

Los videojuegos educativos o también conocidos como “serious games” o juegos serios, son aquellos desarrollados con un propósito formativo o que no tienen como único fin el entretenimiento [Cruz-Lara et al., 2013]. Generalmente se utilizan en ámbitos educativos, científicos o de ingeniería.

Muchos estudios constatan la eficacia de los juegos educativos como método de aprendizaje. El principal motivo es que mientras el usuario juega y se divierte, aumenta la motivación y se elimina casi por completo el estrés. Por otra parte, debemos destacar que en los juegos no suele haber miedo a cometer equivocaciones [Marcano Lárez, 2008] y, por lo tanto, mediante prueba y error, el usuario aprende más rápidamente.

Asimismo, al tratarse de videojuegos, son más llamativos e interactivos para cualquier tipo de público, y consiguen que aumente la concentración del usuario en la tarea que está llevando a cabo. Además, favorecen la competitividad y el trabajo en equipo.

Por último, hay que destacar que no son solo un método de aprendizaje, sino que puede ser combinado con las técnicas tradicionales [Gallego-Durán et al., 2014], haciendo partícipe al docente. Esto se consigue mediante el seguimiento controlado de la actividad que hace el alumno y su posterior análisis por parte del profesor, que puede sugerir mejoras a las actuaciones de este.

#### **2.5.1.2. Classroom VR**

Classroom VR es un entorno virtual diseñado en 2020 por Mario Bocos, Álvaro López y Alejandro Díaz [Bocos Corredor et al., 2020] como Trabajo de Fin de Grado de Ingeniería Informática en la Universidad Complutense de Madrid. Este videojuego trata de mejorar las habilidades de gestión de conflictos de los profesores de secundaria mediante la simulación de distintas situaciones conflictivas en el aula.

Al ejecutar este juego educativo, podrás ver un menú con las tres situaciones de las que está compuesto. Una vez selecciona una de ellas, el juego te situará en un aula de secundaria. Pasados unos segundos, ocurrirá una situación conflictiva. Para darle solución, se presentarán entre tres y cuatro posibles acciones de las que debes escoger una (la más apropiada según el criterio del usuario) y llevarla a cabo. El juego analizará tu posición y algunas palabras clave para determinar qué acción, de las mostradas, has escogido. Para finalizar, se mostrará un feedback sobre la actuación del usuario.

Nuestra aplicación Classroom VR-Motion Capture, ha tomado como base este juego educativo, añadiendo el análisis del lenguaje no verbal (expresión corporal, tonalidad de la voz, proxemia, etc.) y la estimación de emociones según el mismo. Además, hemos mejorado la usabilidad y el diseño del entorno virtual.

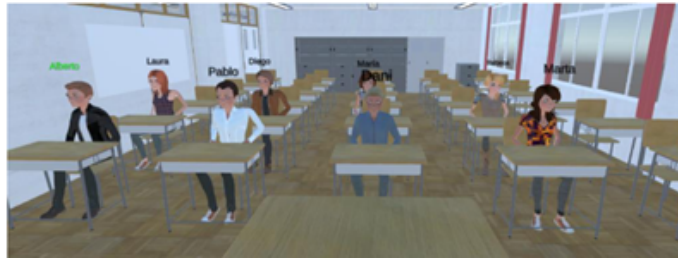


Figura 2.13: Entorno virtual de Classroom VR.

### 2.5.1.3. Didascalía (RTI2018-096401-A-I00)

Didascalía es un proyecto *i+d+i* financiado por el Ministerio de Ciencia, Innovación y Universidades. Su objetivo es crear un entorno virtual basado en realidad virtual que pueda ser utilizado por profesores de educación secundaria y por docentes en formación, por ejemplo, en el Máster de Formación del Profesorado. Además, precisan un análisis del lenguaje no verbal para la mejora de sus capacidades comunicativas al tratar con alumnos.

El equipo Didascalía está formado por profesionales de distintos campos como la psicología, pedagogía, educación e ingeniería. Todos ellos trabajan por un objetivo común, mejorar la formación de los nuevos docentes.

Nuestro proyecto (CVR-MC) se enmarca en Didascalía. Partiendo del proyecto [Classroom VR](#) explicado en el apartado anterior, aportaremos una nueva aproximación al análisis del lenguaje no verbal de los docentes mientras tratan de gestionar el clima del aula. De esta manera, nuestro entorno virtual e investigaciones servirán como base para los próximos años de investigación del equipo Didascalía.

## 2.5.2. La Realidad Virtual

### 2.5.2.1. Qué es

La realidad virtual (RV), o también conocida como ambiente virtual, es considerada una de las nuevas tecnologías en auge de las últimas décadas. Aunque han sido muchos los autores que han intentado definir este concepto, no se ha podido concretar una única definición debido a su carácter abstracto.

La creación de la realidad virtual se remonta a 1965 cuando Ivan Sutherland creó el primer programa en RV para ser utilizado con un casco “inmersivo”. Actualmente entendemos el concepto de realidad virtual según la definición que dio Bernie Roehl en 1996 [Roehl, 1996]: la RV es una simulación tridimensional que es proporcionada por un ordenador, mediante el cual, el usuario es capaz de interactuar libremente con el entorno y sus componentes.

Hoy en día se conocen dos tipos de realidad virtual, la variante no inmersiva o de escritorio que fue la primera en desarrollarse y se caracteriza por la visualización del entorno virtual a través del monitor del ordenador, permitiendo al usuario interactuar con el mismo mediante dispositivos no especializados como pueden ser el ratón, teclado o sensores de posicionamiento. Por otra parte, encontramos la variante inmersiva o de inmersión com-

pleta, es la más conocida por los dispositivos que utiliza el usuario para “adentrarse” en el entorno virtual. Estos suelen ocupar todo el campo de visión y permiten al usuario evadirse completamente del entorno real. Además incluyen mandos o controladores para poder manejar o relacionarse con los elementos del entorno. [Mendoza, 2016] [Escartín, 2000].



Figura 2.14: Visor y controladores Touch de Oculus Rift S

#### 2.5.2.2. Aplicaciones de la Realidad Virtual

La posibilidad de crear ambientes tridimensionales en los que el usuario puede adentrarse hace que esta tecnología sea una buena opción para realizar terapia, entrenamientos o prácticas en entornos seguros.

- **Realidad virtual y medicina**

Son muchos los profesionales que solicitan un cambio en la formación de futuros sanitarios para que esta sea más completa [Vázquez-Mata, 2008]. Aunque la cirugía asistida por ordenador o los simuladores de operaciones quirúrgicas, tienen cada vez mayor peso en la educación de los futuros médicos, aún no están completamente extendidas.

En el campo de la medicina donde el aprendizaje de los alumnos está ligado a posibles negligencias, es necesaria la creación de entornos seguros de práctica con cero riesgo. La realidad virtual es una posible solución para este problema.

La simulación clínica en entornos de realidad virtual permite el desarrollo profesional de los estudiantes mediante la práctica de todas las etapas del tratamiento: diagnóstico, intervención y reflexión. Este tipo de sistemas ha tenido grandes resultados, no solo en la adquisición de nuevos conceptos, sino también en la creación de una sensación de seguridad en los estudiantes que es vital para sus futuras intervenciones reales. [Valencia Castro et al., 2019]

- **Realidad virtual y psicología**

La expansión de este tipo de tecnologías inmersivas ha hecho que sea accesible a multitud de campos, incluido el de la salud. Se han reportado muchos casos de éxito en la aplicación de la realidad virtual para el tratamiento de varios trastornos psicológicos. El éxito de la combinación entre tecnologías inmersivas y terapia se debe a la posibilidad de experimentar situaciones que producen trastornos psicológicos en un entorno seguro, permitiendo pruebas sucesivas y que puede detenerse en cualquier momento para la comodidad del paciente. El primer estudio psicológico realizado con realidad virtual data de 1995 y a este le han seguido multitud de casos.

La exposición de pacientes con trastornos de ansiedad es una de las técnicas más efectivas que toman los psicólogos su tratamiento. Un claro ejemplo son las personas

que sufren fobia social, estas suelen optar por evitar afrontar las situaciones en las que tienen miedo de sentirse juzgados por lo que otros piensen. En este sentido, es difícil crear un entorno seguro real donde el paciente pueda afrontar este tipo de situaciones sin sufrir daños. La solución que provee la realidad virtual mediante la creación de estos entornos ha resultado ser de gran ayuda en este tipo de terapias. [Botella et al., 2006]

Por otra parte, también se ha reportado el éxito del tratamiento de trastornos de déficit de atención con hiperactividad mediante entornos inmersivos de realidad virtual [Delgado Pardo & Moreno García, 2012]. Este tipo de trastornos, que es diagnosticados principalmente en niños, suelen verse ensombrecidos por otros comportamientos que bien pueden estar relacionados con la forma de ser del propio niño. Por este motivo, el diagnóstico en ocasiones se realiza demasiado tarde, habiendo perjudicado el desarrollo personal del menor.

La realidad virtual ha permitido crear entornos de diagnóstico donde se simula un aula y analizan el comportamiento del usuario a distintos estímulos [Rizzo et al., 2002]. Este tipo de herramientas permiten el diagnóstico precoz y en este sentido, proporcionar un tratamiento lo antes posible.



Figura 2.15: Escenas de la clase virtual para el diagnóstico de Déficit de Atención.

#### ■ Realidad virtual y la criminología

En la Universidad de Zúrich, se creó una herramienta Forensic Holodeck [Ebert et al. 2014] capaz de reproducir de manera realista casos criminalísticos para poder ser aportado como prueba en juicios penales. La verosimilitud con la realidad que caracteriza este tipo de herramientas ha permitido esclarecer numerosos casos penales en todo el mundo.

Forensic Holodeck fue utilizado por primera vez para la reconstrucción del escenario donde un policía recibió un disparo en un establecimiento sueco. La exactitud con la que se reprodujo el escenario en realidad virtual permitió que fuera considerada como prueba legal para la propuesta de la condena.

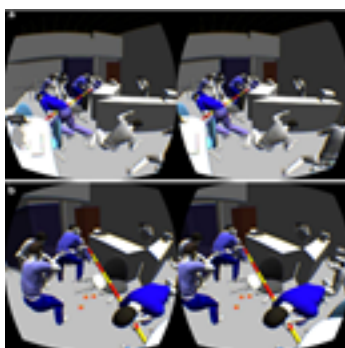


Figura 2.16: Reproducción de la escena del crimen mediante Forensic Holodeck.

Otro caso llamativo donde se empleó este tipo de herramientas fue el caso de Asunta Basterra Porto, el mediático caso español de la niña adoptada que fue asesinada por sus progenitores en 2013. En las sucesivas sesiones del jurado popular, la reproducción del escenario del crimen en realidad virtual fue un factor determinante en la decisión del veredicto. [RAMALLAL & MURILLO, n.d.]

### 2.5.2.3. Realidad virtual y las habilidades de comunicación del profesorado

Contar con docentes cualificados y experimentados al frente de las aulas es imprescindible en cualquier sociedad. Sin embargo, la mayoría de los docentes refieren no haber recibido formación práctica suficiente para gestionar el clima en el aula. La realidad virtual y su capacidad de situar al usuario en cualquier situación y hacer que parezca real, puede proporcionar una solución a este problema.

Este proyecto no es el primer acercamiento a la mejora de competencias a profesores usando entornos virtuales y nuevas tecnologías. En EEUU, crearon el sistema TeachLivETM [Barmaki & Hughes, 2015], una herramienta destinada al análisis del lenguaje no verbal. Mediante un estudio donde solo la mitad de los participantes recibían feedback sobre su lenguaje no verbal y cómo mejorarlo durante sus lecciones, se demuestra la mejora que obtienen aquellos candidatos que obtuvieron feedback a la hora de expresarse corporalmente. Sin embargo, esta herramienta no tenía en cuenta los distintos elementos del lenguaje no verbal como son la mirada o la tonalidad de la voz.

Otro ejemplo fue desarrollado en Japón [Hwang et al.2013], donde crearon un sistema destinado a la práctica docente de profesores noveles. Este entorno simulaba una clase de instituto y contaba con una Kinect con la que se recogía el movimiento de cabeza y manos además de la voz del usuario.

En cuanto a aproximaciones más actuales, tenemos que destacar el estudio realizado por investigadores como Wyss [Wyss et al., 2021]. Desarrollaron un sistema que permite fomentar y mejorar la visión profesional de futuros profesores mediante el seguimiento ocular y verbalizaciones del pensamiento en voz alta post-hoc.

Por otra parte, debemos mencionar el estudio realizado en el Knowledge Media Research Center (Leibniz-Institut für Wissensmedien) ubicado en Alemania, [Goldberg et al., 2021] donde han validado un sistema que permite la evaluación de los procesos relacionados con

la atención de los estudiantes a través de indicadores visibles del grado de participación en el aprendizaje.

Todos estos sistemas aportan información de gran interés sobre cómo afecta la comunicación durante las interacciones con los estudiantes para conseguir las metas educativas. La información recopilada en estos entornos permite evaluar la competencia de los docentes en formación, incluso permite prácticas sucesivas, lo cual es un objetivo formativo fundamental.

#### **2.5.2.4. Ventajas e inconvenientes de la Realidad Virtual**

Por lo general, los diversos campos del conocimiento que incluyen procedimientos basados en la realidad virtual llaman la atención sobre sus numerosas ventajas. Entre ellas destacan las siguientes:

En primer lugar, si comparamos la actitud que tiene el usuario al enfrentarse a una situación real para la que no está preparado, podemos asegurar que se sentirá menos seguro que si, por el contrario, ya ha estado en contacto con este tipo de situación anteriormente. En este sentido, la realidad virtual provee un entorno donde enfrentarse a situaciones reales en entornos virtuales, muy parecidos a la realidad, para que el usuario se sienta más preparado a la hora de enfrentarse a ellos en un entorno real [Botella et al., 2006].

En relación con lo anterior, también debemos destacar que la RV no solo provee de un entorno realista donde poder practicar situaciones desconocidas o complicadas para el usuario, si no que además estos entornos son completamente seguros y permiten múltiples prácticas. En este sentido los usuarios no pueden causar ningún daño mientras realizan las acciones que crean necesarias dentro del entorno, ahorrando en muchas ocasiones los costes propios de las prácticas reales [Cruz et al., 2014].

Debemos destacar, que los entornos virtuales también son aptos para el aprendizaje gradual. Mediante los moduladores podemos ajustar la dificultad de los eventos a afrontar durante la simulación. Esto hace que cualquier usuario, independientemente de sus conocimientos, pueda iniciar su formación con este tipo de sistemas.

Por otra parte, si comparamos la realidad virtual con la imaginación, su comparativa más próxima, la primera resulta mucho más inmersiva y ayuda al usuario a juzgar si realmente está preparado para enfrentarse a esa misma situación en un entorno real.

Por último, este tipo de sistemas inmersivos permite estar acompañado durante la simulación de, por ejemplo, un tutor o profesor [Botella et al., 2006]. De esta manera, no solo el usuario se siente más tranquilo de poder plantear cuestiones durante la situación, sino que además puede desembocar en un debate posterior para la mejora de la gestión de la situación.

En lo referido a los inconvenientes, debemos destacar las limitaciones que supone utilizar este tipo de sistemas. Una de las más importantes es el coste del equipo. Al tratarse de una tecnología inmersiva, son necesarios un visor y controladores, además del software para ejecutar. Por otra parte, no todos los usuarios son aptos para este tipo de simulaciones [SÁNCHEZ-CABRERO et al., 2019] puesto que algunos de estos dispositivos pueden

producir mareos, vértigos, dolores de cabeza, etc. Las personas propensas a este tipo de síntomas no deberían hacer uso de la realidad virtual o prolongar en exceso su uso.

## Capítulo 3

# Metodología y tecnología

### 3.1. Metodología de trabajo

#### 3.1.1. Metodología empleada

Para este proyecto decidimos escoger una metodología ágil ya que cumplíamos todos las características en cuanto al contexto del proyecto: [Montero et al., 2018]

- **Equipos de desarrollo pequeños y ubicados en el mismo lugar:** el equipo consta de cuatro estudiantes de la Facultad de Informática de la Universidad Complutense de Madrid.
- **Los requisitos pueden variar a lo largo del proceso:** al comienzo del proyecto el equipo Didascalia, considerado cliente del proyecto, no sabía exactamente qué funcionalidades quería que implementásemos por lo que la definición de los mismos se hizo a lo largo del mismo proyecto.
- **Pocas restricciones legales:** en calidad de estudiantes y con la finalidad de desarrollar un Trabajo de Fin de Grado, no encontramos restricciones legales.
- **Pocas restricciones en el proceso de desarrollo:** en este sentido, tuvimos la posibilidad de contar con todo el material técnico que necesitáramos y los receptores de CVR-MC, psicólogos y pedagogos del equipo Didascalia, no impusieron ninguna restricción en cuanto al formato.

En lo referido a metodologías ágiles, existen varias. Sin embargo, no todas ellas se adaptan a nuestro tipo de proyecto. Tras una intensa búsqueda sobre la especificación y definición de metodologías, nos decantamos por las mayormente utilizadas: Scrum y Extreme Programming. [Suarez et al., 2016]

Sin embargo, a la hora de decidir cuál de ellas encajaba más con un Trabajo de Fin de Grado basado en la implementación software, revisamos las principales diferencias entre ambas metodologías. Este estudio dejó ver que XP se adaptaba mejor a lo que buscábamos sobre todo en lo referido a las siguiente tabla.



SCRUM	XP
Más basada en la administración del proyecto.	Más basada en la programación del producto.
Iteraciones de 2 a 4 semanas.	Iteraciones de 1 a 3 semanas.
Al finalizar el sprint las tareas aprobadas no se vuelven a tocar en ningún momento.	Todas las tareas son susceptibles de sufrir modificaciones si así lo pide el cliente.
Cada miembro del equipo de desarrollo trabaja individualmente.	Los miembros del equipo de desarrollo trabajan por parejas.

Cuadro 3.1: Principales diferencias entre SCRUM y XP tradicionales.

### 3.1.1.1. Extreme programming (XP)

La metodología eXtreme Programming está enmarcada dentro de las metodologías ágiles. Fue creada por Kent Beck y se considera la primera metodología ágil de la historia. [Amaro Calderón & Valverde Rebaza, 2007].

XP está centrado en las buenas prácticas del desarrollo del software. Esto lo hace promoviendo el aprendizaje de los desarrolladores y creando un buen clima de trabajo donde se promueva el trabajo en equipo. Además, como en la mayoría de metodologías ágiles, se hacen versiones con incrementos muy pequeños haciendo que el código evolucione constantemente.

Entre los valores de esta metodología destacan la simplicidad en las soluciones implementadas, la comunicación fluida entre los componentes del equipo y la retroalimentación por parte del cliente a lo largo de todo el proceso de desarrollo. [Bahit, 2012] y [Muñoz Sanabria, 2013]

### 3.1.2. Equipo XP

Debido a los escasos integrantes del equipo, no se han podido cubrir todos los roles [Amaro Calderón & Valverde Rebaza, 2007] pertenecientes a esta metodología. Sin embargo, los más representativos y relevantes [Duarte & Rojas, 2008] y que por lo tanto hemos dado prioridad son:

- **Desarrolladores:** estiman el coste en tiempo y esfuerzo que supondrá la implementación de los requisitos. Además, definen las características de entrega y el número de iteraciones que serán necesarias. Este papel lo desarrollaron tres componentes del equipo técnico: Andrés Puente, Antonio Luis Solís y Daniel López.
- **Cliente:** describe los requisitos utilizando historias de usuario de manera general y les asigna prioridad. Consideramos el cliente del proyecto al equipo de Didascalia, más concretamente las personas que trabajan en Barcelona. Fueron ellos los que decidieron los requisitos necesarios para obtener un resultado óptimo.
- **Tracker:** encargado de realizar un seguimiento del proyecto. Comunica al resto del equipo el estado del proyecto en comparación con las estimaciones. Determina cuándo será necesario realizar algún cambio para lograr los objetivos de cada iteración. Su

objetivo principal es mejorar las estimaciones futuras. Este rol lo desarrolló Sandra Alonso mediante la supervisión del avance del proyecto.

- **Consultor:** miembro externo del equipo con conocimiento específico en algún tema necesario para el proyecto, en el que puedan surgir problemas. Al tratarse de un Trabajo de Fin de Grado, este papel lo tomaron los tutores: Borja Manero y Alejandro Romero. Sus conocimientos previos del proyecto y de la metodología fueron claves.
- **Coach:** responsable del proceso global. Es el encargado de hacer que todo el equipo conozca y aplique las prácticas XP. Como estudiante del grado de Ingeniería del Software, Sandra Alonso fue la encargada de escoger la metodología y asegurar su cumplimiento.

#### 3.1.2.1. Ciclo de vida XP

El ciclo de vida XP se compone de 6 etapas principales: exploración, planificación de las entregas, iteraciones y fin del proyecto

- **Exploración:** en esta primera etapa, los clientes comunican al equipo las historias de usuario de la primera entrega. A continuación, el equipo de desarrollo comienza a familiarizarse con el entorno y herramientas que se utilizarán para crear la solución.
- **Planificación de entregas:** el cliente establece las prioridades de las historias de usuario que ha propuesto y los programadores estiman el coste en tiempo y esfuerzo que supondrá. Se acuerda un tiempo de entrega que no será mayor a tres meses.
- **Iteraciones:** suponen la parte principal del proceso. Estas están divididas en cuatro subetapas: aprendizaje o familiarización con las nuevas historias de usuario, diseño de la solución, implementación de la misma y testeo o pruebas internas para el correcto funcionamiento. Estas iteraciones duran aproximadamente tres semanas y se llevan a cabo hasta que el cliente ponga fin al proyecto.
- **Fin del proyecto:** el cliente no tiene más historias de usuario o no hay más presupuesto para continuar. En esta fase se genera la documentación final del sistema y no se realizan más cambios en la arquitectura.

#### 3.1.3. Planificación temporal

A continuación, se muestra el diagrama de Gantt que muestra la planificación del proyecto basado en el ciclo de vida XP comentado en el apartado anterior:

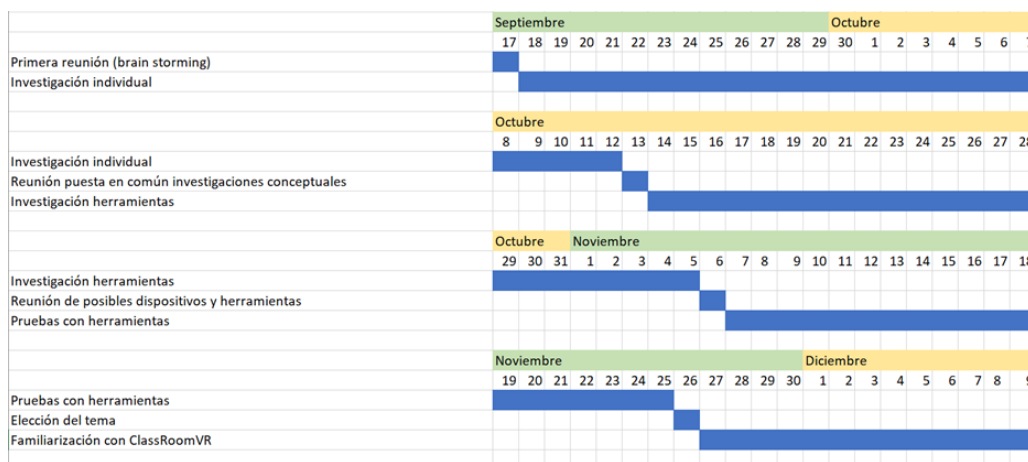


Figura 3.1: Planificación de los meses septiembre, octubre y noviembre



Figura 3.2: Planificación de los meses diciembre y enero

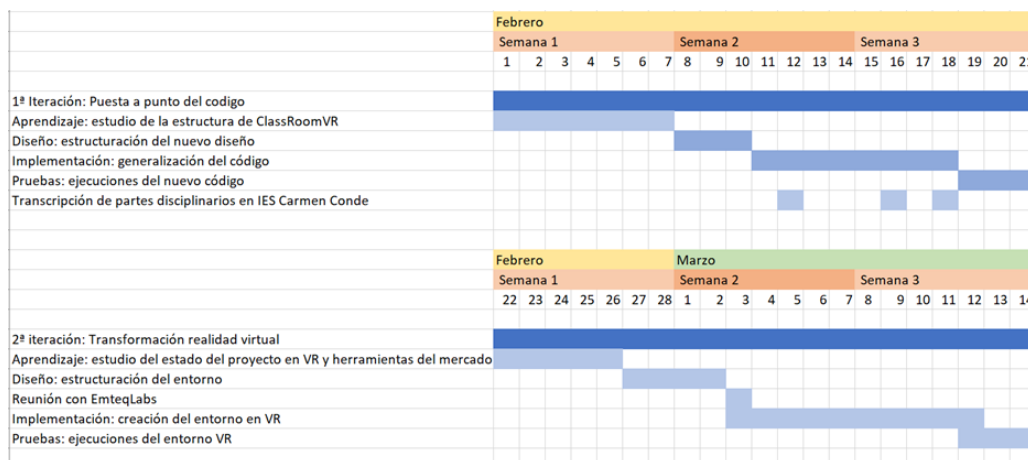


Figura 3.3: Planificación de los meses febrero y marzo

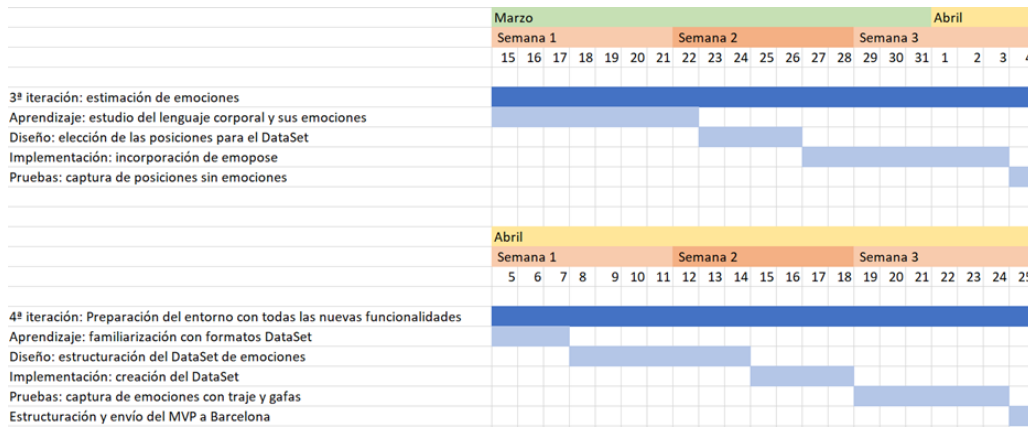


Figura 3.4: Planificación de los meses marzo y abril

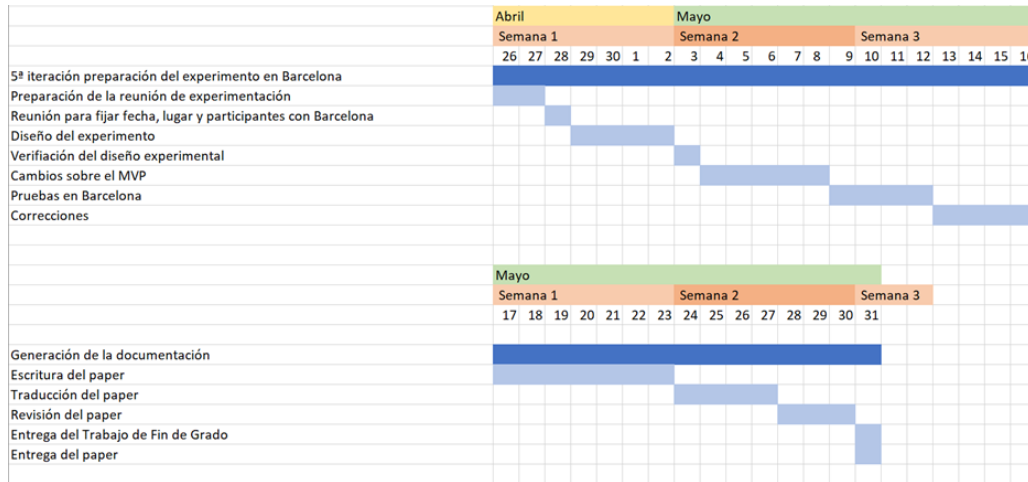


Figura 3.5: Planificación de los meses abril y mayo

### 3.1.4. Control de versiones

Un factor fundamental para el éxito del proyecto es el control de versiones. En este trabajo hemos contado con dos herramientas para organizar el contenido del proyecto y mantenernos al día de las actualizaciones de cada uno de los componentes del equipo. Estas han sido Google Drive y GitHub.

#### 3.1.4.1. GitHub

Para el control de versiones durante el desarrollo del proyecto hemos utilizado Git, el sistema de control de versiones desarrollado por Linus Torvalds con el objetivo de gestionar código fuente. Para el control de repositorio hemos utilizado la plataforma [GitHub](#) más popular para este tipo de trabajo, que aloja y gestiona repositorios Git. Además, esta herramienta nos permite tener un control absoluto de las diferentes versiones contando con un historial donde podemos consultar los cambios realizados en determinada versión o volver a versiones anteriores si fuese necesario.

Para consultar la implementación (código) de Classroom VR-Motion Capture puede entrar en el [repositorio del proyecto](#)

#### 3.1.4.2. Google Drive

La plataforma Google Drive es considerada una de las herramientas de almacenamiento y trabajo colaborativo más versátiles a nivel de educación según trabajos de investigación como [Castellanos Sánchez & Martínez De la Muela, 2013]. La creación de una Unidad Compartida a la que teníamos acceso los cuatro estudiantes y ambos tutores nos permitió hacer un seguimiento de toda la documentación generada durante el desarrollo del proyecto. La estructura de esta unidad está subdividida en diferentes carpetas donde cada miembro tenía los permisos necesarios de lectura y escritura.

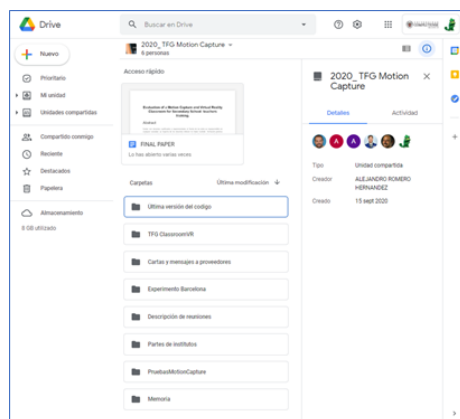


Figura 3.6: Estructura de la Unidad Compartida Drive.

### 3.2. Tecnología base

Existen diversas tecnologías en el mundo del desarrollo de software que nos hubiesen permitido desarrollar la herramienta, sin embargo, sí hay ciertas herramientas que nos facilitan el trabajo ya sea por familiaridad en el uso o su relación con el proyecto.

#### 3.2.1. Unity

Como entorno de desarrollo para nuestro proyecto se eligió continuar con Unity, como hicieron los autores de Classroom VR. Unity es lo que se conoce como un motor de videojuegos, un conjunto de software que permite el diseño, creación y desarrollo de sistemas interactivos (Videojuego).

Este motor nos proporciona un conjunto de herramientas comunes, simulaciones físicas y efectos de iluminación que nos permite desarrollar aplicaciones con facilidad.

Además de ser gratuito, nos permite poder desarrollar hasta en 25 plataformas diferentes, incluyendo algunas de escritorio (Windows, Mac y Linux), consolas de videojuegos (PlayStation, Xbox o Nintendo) y dispositivos móviles (Android y IOS).

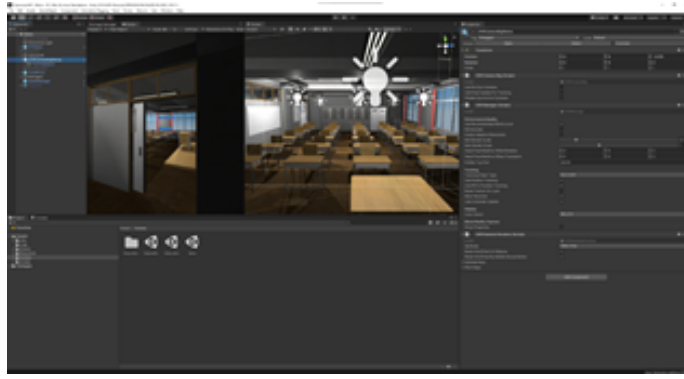


Figura 3.7: Escritorio de Unity

### 3.2.2. C#

Al utilizar Unity hay que pensar en el lenguaje de programación que este requiere, en este caso `C#` que es un lenguaje de programación muy completo y fácil de aprender. Este lenguaje es de tipo *scripting*. Esto hace que los objetos dentro del motor de Unity tomen distintos comportamientos dependiendo del *script* que se le añada, como por ejemplo: establecer la entrada del input, modificar el movimiento de un objeto o gestionar su actividad.

## 3.3. Software utilizado

### 3.3.1. Emopose

A la hora de escoger un sistema que nos permitiera la captura del lenguaje corporal, decidimos escoger [Emopose](#). Entre los motivos principales de esta decisión se encuentran: que es una herramienta sencilla de usar y contiene la funcionalidad de estimación de emociones en base a una posición corporal, un objetivo específico de nuestro proyecto. Además, Emopose está desarrollada en Unity, entorno con el que está familiarizado el equipo de desarrollo.

Por otra parte, debemos destacar que el desarrollador principal de Emopose, Iván García-Magariño, es un profesor de la facultad de Informática de la Universidad Complutense de Madrid, nuestro centro de estudios. Esto nos permitió hacer varias reuniones con él y solucionar algunas dudas que fueron surgiendo a la hora de incorporarlo en CVR-MC.

### 3.3.2. Axis Neuron

La aplicación de Axis Neuron nos permite replicar, en tiempo real, con ayuda de un traje de detección de movimiento que utiliza el usuario, en nuestro caso [Perception Neuron](#), sus movimientos dentro de la aplicación. Gracias a esto se pueden obtener de manera precisa casi todos los movimientos de las diferentes partes del cuerpo.

Axis Neuron permite grabar animaciones, para más adelante, incorporarlas a algún avatar si así se requiere. También permite *stremear* los movimientos capturados por el traje

de detección de movimiento a otra aplicación. Para esto es necesario unas configuraciones previas.

Específicamente en **Unity** lo primero que se deberá hacer es descargar la SDK desde la página web de Perception Neuron e importarla en Unity.

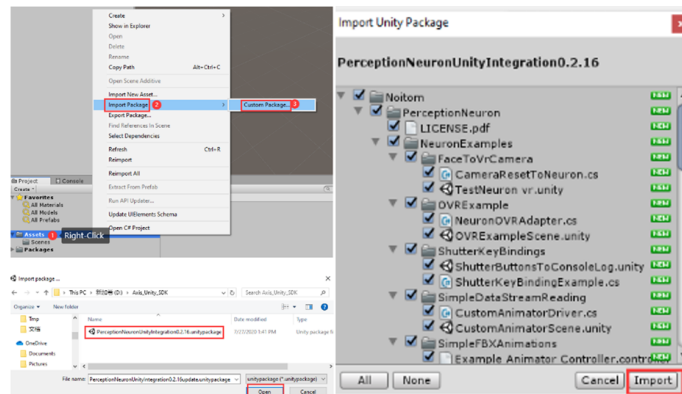


Figura 3.8: Importación del SDK para Perception Neuron en Unity

Por otra parte, el setup de streaming de datos en Axis Neuron se configura de la siguiente manera:

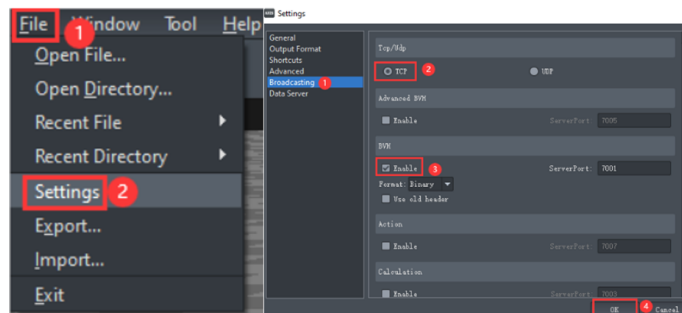


Figura 3.9: Configuración del setup de streaming de datos en Axis Neuron

Una vez realizados estos pasos, ya se estará realizando el intercambio de datos. Podríamos ejecutar una animación grabada previamente o conectar un dispositivo para que capture los datos en tiempo real.

Para ser capaces de recibir este *stream* de datos en Unity tendremos que realizar las siguientes configuraciones:

- Importar PerceptionNeuron.unitypackage
- Importar el modelo del esqueleto que recibirá los datos a Unity.

- En la ventana de proyecto en Unity, hacer *click* en nuestro modelo y seleccionar la casilla de “Rig” en el inspector. Ahí deberemos seleccionar el tipo de animación a “Humanoid” y clickar en apply.
- Si quieres configurar el mapeo de esqueletos, haz click en “configure”, generalmente Unity lo detectara correctamente, aunque no está de más revisarlo por si acaso.
- Confirma los cambios realizados y carga la escena que quieres usar. En la cual deberás incluir el avatar a la jerarquía de la escena.
- Con el avatar en la escena le añadiremos el componente *NeuronTransformInstance.cs*.
- Será importante configurar los ajustes para que sean los mismos que en tu setup (Axis Neuron), activar el “Connect to Axis” permitirá recibir el *stream* de datos. Es importante mantener Axis Neuron activo y ejecutando, con el avatar activado con el traje o la animación en *loop*.

De esta manera al ejecutar la escena de Unity se realizará la acción retransmitida desde Axis Neuron a nuestro personaje de Unity.

### 3.3.3. Software en pruebas

#### 3.3.3.1. MLKit

Como hemos mencionado en el capítulo [Estado del arte](#), MLKit es una solución liviana y versátil para que los desarrolladores de aplicaciones detecten la pose del cuerpo de un sujeto en tiempo real a partir de un vídeo continuo o una imagen estática. Una pose describe la posición del cuerpo en un momento en el tiempo con un conjunto de puntos de referencia esqueléticos. Los puntos de referencia corresponden a diferentes partes del cuerpo como los hombros y las caderas. Las posiciones relativas de los puntos de referencia se pueden utilizar para distinguir una pose de otra.

Aunque esta aplicación nos permite capturar el movimiento del usuario, no incluye la recepción de datos a través de un traje de captura de movimiento. Lo cual buscábamos ya que obtendremos unos resultados más precisos.

#### 3.3.3.2. ThreeDPoseUnityBarracuda

Una herramienta desarrollada en Unity por Yukihiho, [repositorio de GitHub]. Permite por medio de una cámara web proyectar un esqueleto de la persona. La aplicación además, es capaz de utilizar la proyección del esqueleto en un modelo 3D en tiempo real.

Uno de los problemas de esta aplicación es que solo funciona con una única persona, además existe el problema de que la grabación debía ser de cuerpo completo ya que si no la proyección del esqueleto perdía eficacia.



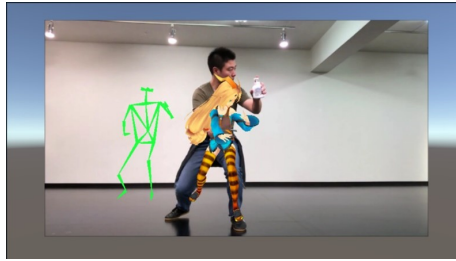


Figura 3.10: Ejemplo de ejecución de la herramienta ThreeDPoseUnityBarracuda

## 3.4. Dispositivos utilizados

### 3.4.1. Oculus Rift S

Las gafas de realidad virtual son dispositivos que permiten visualizar imágenes creadas en el ordenador a una distancia muy cercana de los ojos.

Al igual que otros cascos de realidad virtual, el Rift se coloca como dos lentes en forma de gafas delante de tus ojos, que te brindan una vista estereoscópica en 3D de un mundo en el que estás ubicado. Luego, una combinación de sensores calcula sus movimientos en el mundo real y los traduce en movimientos dentro del juego.

Las Oculus Rift S son un modelo que salió a la venta en la primavera de 2019 por lo que estamos ante uno de los más actuales modelos de gafas de realidad virtual.

Algunas de las características de su Hardware son:

- **Visión:** este dispositivo cuenta con una resolución de 2560 x1440 y una frecuencia de actualización de 80Hz, además de un campo de visión del 115°
- **Audio:** el audio en este dispositivo a diferencia de otros no son auriculares, si no que cuenta con unos pequeños altavoces integrados en el propio casco.
- **Oculus Insight:** este nombre es el que se le ha dado al sistema de seguimiento de 5 cámaras que componen el dispositivo. Estas rastrean movimiento mediante diodos de infrarrojos, además de acelerómetros y un motor de predicción.
- **Controladores:** utiliza los mismo controladores que la anterior generación con la inclusión de los anillos superiores para ser detectados por los sensores.



Figura 3.11: Componentes de Oculus Rift S

### 3.4.2. Perception Neuron

[Perception Neuron](#) es un traje de captura de movimiento desarrollado por [Noitom](#), una empresa dedicada al desarrollo de tecnologías para la captura de movimiento.

El traje consta de un conjunto de correas con velcro que se colocan en las zonas del cuerpo que quieren ser analizadas: cabeza, torso, brazo (antebrazo y hombro), manos, cintura, piernas (parte superior e inferior) y pies. Todas ellas con neuronas que van conectadas unas a otras hasta el conector general, situado en la cintura, que es el que ha de ser conectado al equipo en cuestión. Una vez conectado, envía las métricas extraídas del movimiento de las neuronas.

Este sistema ha demostrado ser de utilidad para muchos estudios entre ellos el llevado a cabo por Matías Macías y Simón Pineda de la Universidad Pontificia Bolivariana [Macías Gómez & Pineda Agudelo, 2020] llamado *Generación y control de contenido Audiovisual utilizando TouchDesigner, Ableton Live y los movimientos del cuerpo a través de Perception Neuron*. Con este trabajo pudieron utilizar este traje para introducir al usuario en un entorno virtual otorgando funcionalidad al movimiento realizado, como producir, recortar y editar sonidos.



Figura 3.12: Componentes del dispositivo Perception Neuron.

### 3.4.3. Ordenador para ejecutar el programa

Para ejecutar la aplicación es necesario contar con un equipo ya sea sobremesa o portátil. Este debe contar con USB 3.0 o bien DisplayPort o mini DisplayPort. Este requisito viene dado principalmente, por la necesidad de conectar las gafas de realidad virtual Oculus Rift S.

Por otra parte, es necesario tener en cuenta la compatibilidad de la tarjeta gráfica para poder ejecutar la herramienta. Esta debe ser una de la siguiente [lista de tarjetas gráficas], donde podemos encontrar desde una GTX 780 Ti hasta la nueva generación RTX.

### 3.4.4. Dispositivos en fase de pruebas

#### 3.4.4.1. Kinect

Tras la investigación previa sobre las herramientas existentes en la captura del movimiento, barajamos distintas opciones y el dispositivo Kinect fue una de las más apoyadas para ser utilizada en el desarrollo del proyecto. Revisando varios proyectos que empleaban este dispositivo como por ejemplo [Machine Learning Movement], un proyecto realizado por Kat Sullyvan, vimos que era posible la clasificación de determinadas posiciones como se muestra a continuación.



Figura 3.13: Pruebas para la lectura del lenguaje corporal realizadas con Kinect

Sin embargo, observamos que el usuario debía estar situado en un fondo liso y mantener aproximadamente 2 segundos la posición que quería ser reconocida. Además, es necesario mucho espacio para que la Kinect pueda enfocar el cuerpo entero del usuario y sus movimientos por la sala.

Otro dato a destacar es que las posiciones no podían superponer elementos del cuerpo, lo que impedía hacer posiciones de enfado como por ejemplo cruzar los brazos o piernas.

Todos esta información y pruebas hicieron que descartásemos Kinect como dispositivo para la captura y análisis del lenguaje corporal.

#### 3.4.4.2. Gafas Emteq Labs

[Emteq Labs](#) es una empresa que trabaja en el ámbito de la realidad virtual, centrándose en la tarea de monitorizar expresiones humanas e interpretarlos por medio de sus dispositivos. Su principal herramienta, **EmteqPRO**, se trata de una serie de sensores musculares que van directamente integrados en las gafas de realidad virtual los cuales miden de manera

biométrica la actividad de la cara del individuo con el fin de entender el comportamiento del usuario



Figura 3.14: Gafas EmteqPRO

Aunque lo consideramos un producto factible y nos reunimos con los desarrolladores, no fue posible utilizarlos finalmente en el desarrollo. Durante la reunión nos confirmaron que se trata de un proyecto aún en fase de pruebas y debido a su fecha de salida al mercado, era muy difícil compaginar su integración en el proyecto a tiempo para la fase experimental, por lo que decidimos no continuar con esta propuesta. Sin embargo, mantendremos el contacto con este proveedor porque sí podrán ser integradas en este proyecto en futuras futuras fases o ampliaciones.

## Capítulo 4

# Diseño y desarrollo de la herramienta

### 4.1. Diseño del proyecto

#### 4.1.1. Fase divergente

El primer paso a la hora de empezar un nuevo proyecto siempre es decidir el tema a tratar. En nuestro caso, aunque la propuesta de Trabajo de Fin de Grado era bastante clara: *Análisis de la expresión corporal utilizando tecnologías de captura de movimiento*, no tardamos en darnos cuenta de la gran cantidad de aplicaciones que podía tener esta tecnología. Por este motivo, las primeras reuniones realizadas en octubre de 2020 fueron dedicadas a hacer lluvia de ideas sobre los diferentes temas que nos gustaría tratar. Algunos de ellos fueron:

- Herramientas para mejorar la expresión corporal de los profesores a la hora de impartir clases.
- Herramientas para la mejora de posiciones en deportes como la halterofilia.
- Herramienta para comprobar si los niños entienden lo que el profesor explica y si están siguiendo la clase.
- Herramienta para el diagnóstico precoz del Parkinson.
- API para calcular la mejor ruta para escalar un rocódromo.
- Herramienta para detectar el nerviosismo en personas.

Todas estas ideas tienen un factor común: el análisis corporal del usuario con dispositivos de captura de movimiento. Tras varias semanas, tomamos la decisión de escoger la primera de ellas: **“Herramientas para mejorar la expresión corporal de los profesores a la hora de impartir clases”**. Rápidamente vimos la clara relación que tenía el tema con el proyecto Didascalia del que ya nos habían hablado Borja Manero y Alejandro Romero, nuestros tutores.

#### 4.1.2. Incorporación a Didascalía

Una vez decidida la dirección que tomaría del proyecto, Borja Manero y Alejandro Romero propusieron presentarnos como parte del equipo técnico al resto de integrantes del equipo [Didascalía](#). Las primera reuniones fueron algo menos técnicas con el fin principal de empaparnos de todo el entorno en el que empezaríamos a trabajar próximamente.

En las siguientes reuniones, comenzamos a analizar los requisitos que solicitaban los docentes, pedagogos y psicólogos del equipo. Enseguida nos dimos cuenta de que, debido a la situación sanitaria del año 2020, los estudiantes que desarrollaron [Classroom VR](#) no pudieron trasladarse a Barcelona a realizar los experimentos de dicha aplicación. Es por este motivo por el que no todo el equipo Didascalía había visto ni probado nunca la aplicación. Para que todos partiéramos desde el mismo punto y con toda la información presente, hicimos algunos vídeos de presentación que mostraban la ejecución de las escenas de Classroom VR [Bocos Corredor et al., 2020].

Más tarde comenzó la etapa de investigación. Fue clave para ambas partes del equipo, el intercambio de información, documentos y conocimientos. Pudimos comprobar en todo momento que trabajar en un equipo multidisciplinar te hace adquirir rápidamente conceptos sobre distintas áreas.

A partir de este momento comenzó el diseño de requisitos con reuniones semanales para aprobar cambios y mejorar las descripciones de los mismos. Fuimos depurando la aplicación hasta obtener un diseño inicial.

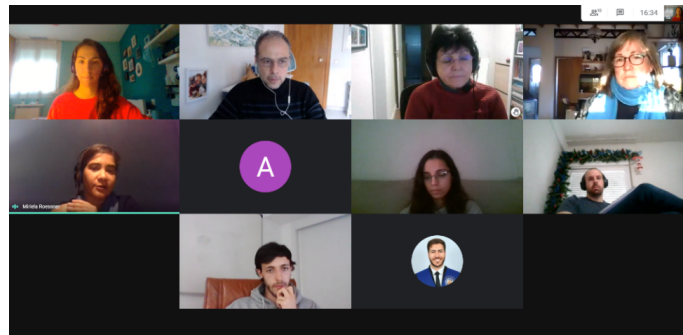


Figura 4.1: Primera reunión con el equipo Didascalía (17/12/2020).

#### 4.1.3. Diseño inicial

El primer paso que llevamos a cabo fue una investigación inicial. Estudiando y obteniendo diferentes técnicas, [herramientas y aplicaciones que permitieran la captura de movimiento](#) y el estudio del lenguaje corporal asociado a emociones. Seguidamente, iniciamos unas pequeñas pruebas con distintos dispositivos y software como [Kinect](#), [Axis Neuron](#) y [Oculus Rift S](#) entre otros.

El proyecto base del que partimos fue el desarrollado por Mario Bocos, Álvaro López y Alejandro Díaz bajo el nombre de *Classroom VR: a VR game to improve communication skills in secondary-school teachers* [Bocos Corredor et al., 2020].

La aplicación consistía en un entorno ambientado en un aula de instituto con alumnos sentados en sus pupitres. Ante el usuario se presentan 3 escenarios disponibles. En todos ellos se seguía la misma estructura: primeramente, hay un tiempo de inactividad en el sistema hasta que aparece una situación disruptiva en el aula. En este momento el usuario (docente novel) puede ver en pantalla un menú con las posibles opciones que tiene para solventar el conflicto. El usuario realiza la acción escogida y el sistema interpreta el camino escogido mediante la posición del usuario con respecto a los alumnos virtuales, determinadas palabras clave que ha pronunciado y la tonalidad de la voz empleada. Una vez acabada la simulación se proporciona un feedback visual de su desempeño ante cada situación.

Además, estuvimos en contacto con el equipo de Didascalia, como receptores del producto, para realizar con ellos parte del análisis de objetivos y las metodologías docentes involucradas.

Paralelamente, nos pusimos en contacto con Iván García-Magariño, desarrollador de [hyperref\[sub:emoPose\]](#)[Emopose](#) [García-Magariño et al., 2019]. Tras evaluar las posibilidades del proyecto, y habiendo decidido incluir la captura de movimiento a la aplicación, tomamos la decisión de incluir su trabajo adaptándolo a nuestras necesidades para evaluar el movimiento transmitido por las tecnologías de Motion Capture en lugar del maniquí de Emopose.

Una de las ampliaciones que pensamos en añadir al proyecto inicialmente fue la utilización del dispositivo de [Emteq Labs](#) llamado [EmteqPRO](#). Se trata de unos sensores que son añadidos a las gafas de realidad virtual y que permiten medir la actividad biométrica de los músculos faciales, permitiendo el análisis de la expresión facial. No obstante, se trata de un proyecto aún en fase de desarrollo y debido a su fecha de salida era muy difícil compaginar su integración en el proyecto a tiempo para la fase experimental, por lo que se ha quedado la idea para una siguiente fase.

#### 4.1.4. Transcripción de partes disciplinarios

Con la finalidad de crear una serie de escenarios realistas y de utilidad para el profesorado en formación, es imprescindible contar con información de qué conflictos son los más frecuentes en el aula y cómo se desarrollan. Para ello, hemos contactado con el I.E.S Carmen Conde, informándoles acerca de nuestro proyecto y nuestra necesidad a este respecto. Ellos nos han permitido acceder a los partes de disciplina de los últimos años.

A lo largo de 3 días, nos fuimos turnando en grupos de 2 para llevar a cabo la transcripción, en la que revisamos más de 600 partes de disciplina. Los campos apuntados fueron los siguientes:

- Curso y grupo del afectado.
- Fecha y hora del incidente.
- Asignatura durante cuya clase ocurrieron los hechos.
- Tipo de incidencia: Leve o Grave.
- Contexto y explicación del suceso.

- Medidas tomadas por parte del profesorado y el centro.

## 4.2. Desarrollo de la herramienta

### 4.2.1. Reestructuración de la aplicación

Se ha reestructurado el código de la aplicación Classroom VR para tratar de abstraer y generalizarlo, de manera que se puedan incluir situaciones críticas fácilmente en el futuro, con sus respectivos caminos y reacciones.

#### 4.2.1.1. Estructura del proyecto

El proyecto se sostiene bajo la siguiente estructura de escenas:

- **Menú:** es la escena encargada de los menús iniciales del juego y la primera en instanciar el *Game Manager*, este objeto se encargará de gestionar el estado general de la aplicación, enviar información entre los diferentes elementos de la misma y la gestión del progreso.
- **Class Game Scene :** es la escena principal de juego en la aplicación, es la escena de la clase donde se desarrollaran las diferentes situaciones. Cuenta con diferentes objetos “*canvas* y *player* en función de la versión del juego que se esté utilizando (VR vs NO VR). También incluye un objeto *My Scene Manager* encargado de gestionar la lógica de la escena, se entrará en más detalle sobre el *MyScene Manager* más adelante.
- **Capture Pose:** es la ecena que implementa el código de [Emopose](#) cedido por Iván García-Magariño para añadir posturas a la base de casos de cada una de las emociones que se analizan.

#### 4.2.1.2. Estructura y generación de las situaciones

Pretendemos que las situaciones que se implementan en la aplicación sean lo más fácilmente configurables y extensibles posible. Para ello, el *GameManager* que se encuentra en la escena *Menu*, recibe como parámetro desde el editor un array de *Scriptable Objects* de tipo *Scene Package*. Con esta información, el *GameManager*, creará todos los botones que sean necesarios para poder seleccionar las distintas escenas. Una vez elegida una situación, cambiará a la próxima escena de la aplicación, en esa escena, le prestará la información del *Scriptable Object-Scene Package* elegido por el usuario, al *MyScene Manager* y este segundo se encargará de generar la escena seleccionada.

Gracias a esta información se tienen los parámetros necesarios para generar una situación. La estructura de los paquetes *Scene Package* es la siguiente:

- **Información para la generación de la clase:**
  - **NStudents:** número de estudiantes en la situación.
  - **Problematic Students:** número de estudiantes problemáticos.



- ***Problematic Together:*** booleano que indicará si los estudiantes problemáticos se generarán juntos o no.
  - ***NGroups:*** número de grupos que la clase generará.
- **Recursos específicos de la situación :**
- ***Ini Message:*** mensaje inicial con la información de la situación.
  - ***Time Start:*** tiempo para que empiece a darse la situación.
  - ***Problematic Animation:*** animación por defecto para la situación crítica
  - ***AudioSituacionMasculino:*** audio por defecto de hombre.
  - ***AudioSituacionFemenino:*** audio por defecto de mujer.
  - ***Audio Reacción Clase:*** audio por defecto de la reacción de la clase a la situación.
  - ***Specific Behaviour:*** en caso de que la información anterior no sea suficiente para realizar la situación prevista, se ha añadido la posibilidad de incluir un método implementado en otra clase que dejará al programador tener vía libre para ejecutar la situación como se desee. Este método deberá comprobar que la situación se ha llevado a cabo correctamente y en ese momento devolver *true*.
- ***Paths:*** array de *Scriptable Objects-Path Package*. Cada uno de estos objetos representa un posible camino a tomar en esta situación. La estructura de estos paquetes es la siguiente:
- ***PathInfo:*** información para que el usuario sepa cómo activar este camino.
  - ***KeyWords:*** array de palabras claves que se analizarán para elegir este camino.
  - ***GetClose:*** booleano para indicar que este camino también se activará si te acercas al estudiante problemático.
  - ***Ignore:*** booleano que indica que este camino es de ignorar la situación
  - ***Audio:*** audio de reacción de la clase al camino.
  - ***Path Class Animation:*** animación de reacción de la clase al camino.
  - ***Path Prob Animation:*** animación de reacción del estudiante problemático. Si no existe ejecutará la misma que el resto de la clase.
  - ***Specific Behaviour:*** al igual que con los *Scene Package* existe la posibilidad de gestionar más la reacción específica de los alumnos a un camino, pasándole un método con el mismo comportamiento que el *Specific Behaviour* the *Scene Package*, es decir que al terminar de ejecutarse devuelve *true* y en caso contrario se le llamara recursivamente.
  - ***Feedback Path:*** *feedback* final al camino tomado.
  - ***Correct Path:*** boolean para indicar que este camino se considera una reacción positiva a la hora de gestionar la situación presentada en la clase.

Es importante indicar que la mayoría, sino todos, los audios, animaciones y *specific Behaviours* que se encuentran en estos paquetes, no es necesario añadirlos, en caso de que algunas escenas o caminos no los requieran, y la escena se seguirá ejecutando correctamente.

Sin embargo, las informaciones previas que indican la situación, la información de cómo tomar el camino y el feedback al mismo si es importante añadirlas.

#### 4.2.1.3. Escena Menú

La escena del menú es la que se presenta al iniciar la aplicación. En ella encontramos un pequeño menú donde podremos navegar y seleccionar de entre las distintas situaciones a resolver. Los objetos que se pueden encontrar en esta escena son:

- **Canvas:** el *Canvas* en Unity es un *GameObject* contenedor de todos los elementos del *UI*. Normalmente el canvas es representado por un área con forma de rectángulo, donde todos los objetos hijos pertenecen a esta área, sin embargo, en la escena se diferencian dos *canvas*, uno para realidad virtual y otro para cuando no está activa la plataforma.
- **Managers:** En esta escena contamos con *GameManager* y el *UiManager* encargados de la visualización y funcionamiento de la escena.

#### 4.2.1.4. Escena Class\_GameScene

La escena *Class\_GameScene* es la principal y donde se desarrollan las situaciones de la herramienta. Los objetos que podemos encontrar en la escena son:

- **Canvas:** se diferencian el *Canvas* para la plataforma de PC y el *CanvasVR* para realidad virtual. Se activa o desactiva el indicado en función de la plataforma de juego que indique el *GameManager*.
- **Scene Objects:** objeto vacío que es padre de todos los objetos que se encuentran inicialmente en la escena de juego.
- **Managers:** los diferentes objetos, de los managers presentes en la escena, gestionan el correcto funcionamiento de la misma, todos ellos centralizados en el *SceneManager* y este, a su vez, en el *GameManager*.



Figura 4.2: Estructura de Managers de CVR-MC.

- ***SeceneManager:***

Cuenta con un único componente que le ayuda a realizar todas las acciones de la escena:

- ***MySceneManager:*** clase principal que se encarga de la ejecución de la escena. Cuenta con un booleano, para poder confirmar que la generación de la escena se ha realizado correctamente, que hace que la escena se ejecute o no. Tiene referencias a: *EmoPose UiManager*, *Player Motion*, *Scene Objects* y *Sound Manager*.

Esta clase cuenta con un sistema de “Máquina de estados” para gestionar el estado del juego en función del momento en el que nos encontremos. Se puede dividir la ejecución de la escena en 3 estados principales que a su vez se dividen en subestados.

- ◊ **Presentación de la situación:** (Inicio de la escena- fin de la presentación de la situación)
- ◊ **Toma de decisión del camino a tomar:** (fin de la presentación de la situación - toma de decisión del usuario)
- ◊ **Fin de resolución de la situación:** (Decisión del camino por el usuario - fin de la ejecución de la acción escogida por el usuario)

Los estados secundarios de la escena son: *AnimSituation*, *AnimReactSituation*, *GeneratePathSettings*, *ChoosingPath*, *ReactToPath*, *ShowFeedback*. Estos estados secundarios hacen referencia, de dos en dos, a los estados principales.

Tras la explicación de los *scriptable Objects* y la estructura de *My Scene Manager*, es fácil imaginarse cómo funciona por detrás la lógica de *MyScene Manager*, que va analizando la información que ha recibido para ir desarrollando la escena. Además, *MySceneManager* también recibe un *ScriptableObject-ClassInfo*, recibido gracias al *GameManager*, que contiene la información del prefab de la clase, los nombres de alumnos masculino-s/femeninos que se usarán los prefabs de los alumnos, su animación *Idle* y el *student Animator controller* que usarán los alumnos para ejecutar todas sus animaciones.

- ***MotionCapture Manager:***

Este objeto manager cuenta con 3 componentes que le permiten realizar todas las actividades relacionadas con la captura de movimiento.

- ***Motion Capture Manager:*** se encarga de todo lo relacionado con la captura de la emoción en función de la posición del cuerpo del usuario. Además, guarda esa información para mostrarla más adelante al final de la ejecución de la simulación y guardarla en formato csv en el archivo correspondiente a la situación que se esté realizando.

El *mySceneManager* va llamando al update del *MotionCaptureManager*, para que este capture la emoción cada cierto tiempo. Tras el trabajo de comprensión de *Emopose*, se llevó a cabo una modificación en el mismo, para que, al guardar las posiciones de las extremidades del usuario, no se tuviera en cuenta la posición global del avatar, sino únicamente las posiciones locales.

- ***Pose Builder:*** recibe los *transforms* necesarios para poder generar una pose en función de las posiciones de las extremidades del avatar. Más ade-

lante utilizará esa pose para compararla con el resto de las ya guardadas por el entrenamiento y asociarla a una emoción. Estas extremidades que recibe son: *Left hand*, *Right hand*, *Left foot*, *Right foot*, *Head*, *Look Pos*, *Opening hands Handler*.

- ***Opening hands Handler***: esta clase se encarga de gestionar la apertura de las manos que produce el usuario. No ha llegado a utilizarse, ya que los usuarios contaban en todo momento con unos mandos de Oculus, por lo que la información que nos proporcionaba no iba a ser relevante. Sin embargo, pese no haberse utilizado, EmoPose requiere que lo tuvieras creado para las referencias que utiliza.

- ***Sound Manager***:

El objeto *SoundManager* almacena tanto los componentes como las referencias en código de *MicToVokaturi*, *SoundLoudness*, *KeywordRecognizer* y *MicrophoneManager*. Es el encargado de inicializar la instancia de *KeywordRecognizer* en el método *Awake* de Unity y de escribir la serialización de la información aportada por *Vokaturi* en el método *processVokaturiInfo*.

- ***MicToVokaturi***: es un script a modo de *Wrapper* de *Vokaturi*, que analiza el audio recogido y clasifica las muestras en 5 emociones distintas en su versión gratuita, felicidad, tristeza, neutral, ira y miedo.
- ***SoundLoudness***: recoge muestras de sonido del micrófono mediante la API de Unity y analiza si ha habido un incremento o descenso del tono de voz, lo que permite saber si el usuario ha reaccionado de manera excesiva en respuesta a ciertos eventos que se dan durante el escenario, comparándolo con su propio tono medio dentro de cada escenario.
- ***KeywordRecognizer***: es una clase de Unity Core que es utilizada para identificar palabras predefinidas provenientes de una fuente de audio. La clase contiene un diccionario que debe ser rellenado previamente, que permite lanzar eventos según la palabra reconocida. Este script nos permite controlar el flujo del programa según actúe el usuario.
- ***MicrophoneManager***: es la parte encargada de detectar el micrófono predefinido gracias a la API de Unity, y asignarlo como fuente de audio para el resto de scripts que lo necesiten, así como activarlo o desactivarlo. Por defecto, el micrófono predefinido se asigna en la posición 0, en el valor *deviceNumber* y guarda el nombre del dispositivo en el campo *\_device*.

- ***UiManager***:

El *UiManager* es un objeto que se compone del script *UiManager* que se encuentra en las escenas *Menú* y *GameScene*, su principal función es controlar el canvas y su interacción con el mismo.

Podemos dividir su funcionamiento en dos secciones. La primera se encarga del funcionamiento de los paneles del canvas durante todo el proceso, mostrando u ocultando paneles y textos a lo largo de la escena. La segunda se encarga de gestionar la interfaz en función de si se está utilizando la realidad virtual o no, activando o desactivando distintos objetos necesarios para cada una de las plataformas.

- ***GameManager***:

El objeto *GameManager* es un objeto *dontDestroyOnLoad*, que se mantiene a lo

largo de las escenas. Sin embargo, cada una de las escenas (*Menu* y *Class.GameScene*) cuentan con su propio objeto *GameManager* con la información que requiere este objeto de esa escena, así cuando se pasa de una escena a otra, el *GameManager* obtiene la información del otro objeto de su mismo nombre y tipo ya creado en la escena y luego se destruye la copia de este. Quedando un único *GameManager* con toda la información.

Las referencias que tiene este objeto y la escena en cuestión donde se le indican son:

- ***SceneManager***: referencia al objeto con el *script SceneManager* (escena *Class.GameScene*)
- ***Packages***: un array que tiene todos los *scriptable Objects* de las diferentes situaciones que puede generar la herramienta (escena *Menu*).
- ***ClassInfo***: con la información de la clase (escena *Menu*). Tanto los paquetes de las diferentes escenas a simular como el *class info* se pretende que se rellenen en la escena *Menu*, sin embargo, por comodidad durante el desarrollo, para evitar estar siempre pasando de una escena a otra para hacer pruebas seleccionando la situación que queremos simular. La escena *Class.GameScene* tiene estos dos parámetros también rellenos, aunque solo con una situación para que ejecute esa.
- ***VR Hardware***: este flag nos permite indicar al *GameManager* si el juego se ejecutará con el hardware de VR, y por tanto, todos los menús y lógica deben ser en formato VR o no, en función de lo que se indique.

Los diferentes managers que podemos encontrar a lo largo de la herramienta suelen siempre acabar como punto en común en el *GameManager*, de una manera u otra.

Por ejemplo, al iniciar la aplicación de manera normal, desde la escena *Menu*, y seleccionar la situación que se quiere simular. El *GameManager* guarda esa información que le proporciona el *MenuManager* de la escena *Menu*, más adelante, al cambiar de escena, le proporciona la información de la situación elegida al *SceneManager*, que procederá a generarla y ejecutarla. Una vez terminada la situación, el *SceneManager* informa al *GameManager* del desarrollo final de la situación, para que, si el usuario ha elegido un buen proceso de resolución de la situación, esta se pueda marcar como completada en los menús si así se deseara. Actualmente no está implementado que se marquen las situaciones como completadas.

- ***CSV Serializer***:

Este *script* no es necesario incluirlo como componente en ningún objeto. Al tratarse de una clase estática, permitimos a cualquier *manager* de la herramienta llamar a sus métodos en cualquier momento para guardar la información que se le pase en el archivo correspondiente a la situación que se está ejecutando en ese momento.

El serializador de los datos se inicia al comenzar cualquier escena, recibiendo el nombre de la escena que se está ejecutando. Primero comprueba que el directorio donde va a guardar el archivo existe, en caso de no existir, lo crea. A continuación, busca en ese directorio si existen archivos con el mismo nombre que la escena que se está ejecutando, ya que a cada archivo lo guarda con el

nombre de la escena, seguido del número correspondiente al número de archivos con ese nombre en el directorio.

#### 4.2.1.5. Escena CapturePose

Esta escena se utiliza para entrenar el sistema de [Emopose](#) y guardar posiciones asociadas a emociones en la base de datos de [Emopose](#). Cuenta con un avatar inicial al cual se le ha integrado el componente de [Axis Neuron Neuron Transform Instance](#) para permitirle recibir el *streaming* de datos que Axis proporciona. De esta manera podemos colocar el avatar en la posición exacta que queremos asociar a una emoción.

Previamente se contaba con la escena proporcionada por el proyecto de Emopose, en la cual podías mover al avatar únicamente en algunos de sus huesos.

Una vez que se ha iniciado la escena, y te has posicionado como quieres, se le pide al sistema que compare la posición actual del avatar con la base de datos y calcule la emoción más cercana a esa pose, a través del botón *DONE*. Al mostrarnos la emoción más cercana a esa pose podremos querer guardarla como una posición nueva, asociada a una emoción diferente (o la misma), para ello le daremos a *DISCARD*, lo que nos permitirá seleccionar la emoción a la cual se asocia esa pose e incluirla en la base de casos.

Siempre podemos cancelar el guardado de la nueva pose capturada a través del botón *DISCARD* o guardarla con *SAVE*.

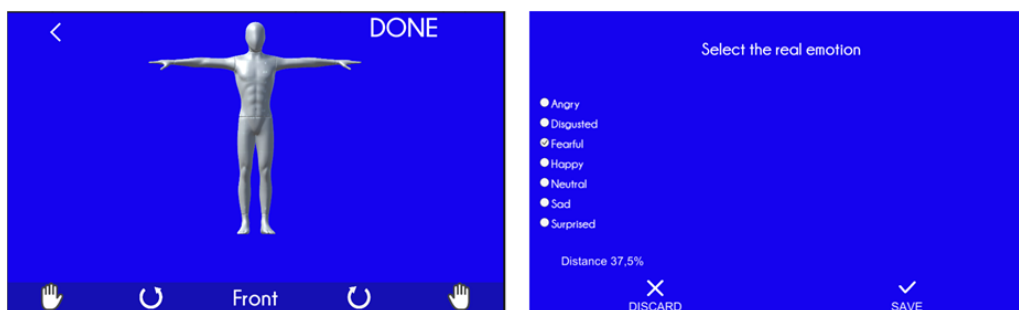


Figura 4.3: Captura de nuevas posiciones mediante la interfaz de EmoPose.

#### 4.2.2. Implementación de la aplicación en Realidad Virtual

Inicialmente este proyecto estaba pensado para utilizarse con realidad virtual, pero aún no se había implementado. La solución que propusimos fue desarrollar un sistema que permitiese cambiar entre ambas plataformas para que pudiera ser utilizada con o sin realidad virtual.

En el proceso de integración debíamos decantarnos por alguna librería que nos permitiera aprovechar las funcionalidades de las Oculus en [Unity](#). Oculus cuenta con integración a múltiples plataformas como *Unreal Engine*, *Native Android*, *Native Window* y *Unity*, que es nuestro caso. Finalmente nos decantamos por utilizar el propio software de [Oculus](#) (Unity Integration) Este es un paquete que se importa en [Unity](#) y nos da la posibilidad de utilizar Scripts prefabricados, muestras de ejemplo y muchos más complementos para desarrollar nuestra aplicación en Unity.

Tras descargar el paquete desde el [Asset Store] e importarlo a nuestro proyecto, como se muestra en la figura 4.4, comenzamos con el desarrollo y transformación a realidad virtual. El primer paso fue activar la realidad virtual desde los ajustes del propio [Unity](#) para que al iniciar la aplicación se pudiese ejecutar en realidad virtual.

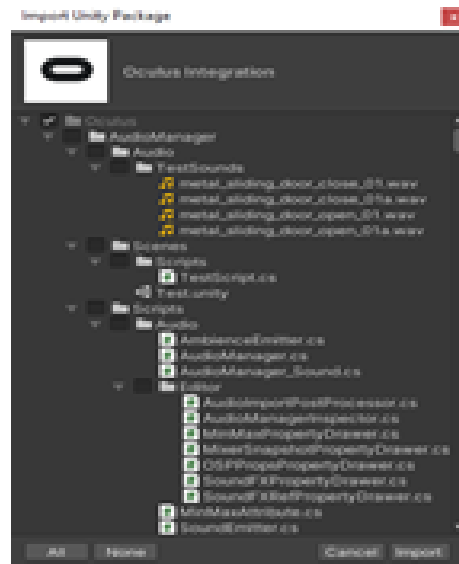


Figura 4.4: Importación al proyecto

Utilizando los componentes que nos brinda la integración de [Oculus](#) creamos nuestro GameObject Player VR que es el objeto del profesor en nuestra clase. Este contiene varios scripts como el movimiento controlado por los mandos de las oculus o el de la cámara. Además, a este player se le añaden las manos que representan la proyección de los controladores en la aplicación.

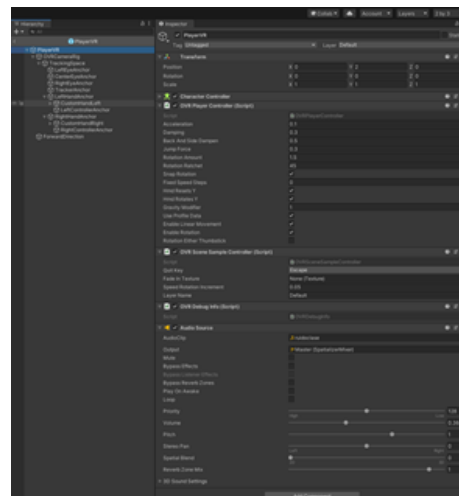


Figura 4.5: Estructura del proyecto

Respecto a la transformación de la interfaz visual, se creó un nuevo canvas (Canvas VR) que está sujeto a la cámara, esto hace que dependiendo de dónde mire el usuario en realidad virtual, toda la interfaz incluida en el canvas le siga.

Para poder interactuar con el nuevo canvas fue necesario incluir UiHelppers, que nos permiten la interacción de los controladores de las [Oculus](#) por medio de láseres que se despliegan de las manos previamente incluidas en el Player VR.

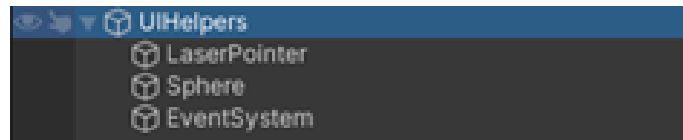


Figura 4.6: Despliegue de UiHelppers

#### 4.2.2.1. Lógica y animaciones de los estudiantes

Respecto a los cambios realizados en la lógica de las animaciones de los estudiantes, al tener una estructura de escenas generales que se dividen en 3 estados (Presentación, Decisión, Reacción a la decisión) y unos comportamientos de alumnos similares en muchos puntos de estos estados. Se decidió implementar por código los comportamientos generales, como ruidos o animaciones, en los momentos determinados de cada estado. Dejando además, la posibilidad al programador de implementar comportamientos específicos dentro de cada estado en un script independiente.

Para la gestión de las posibles animaciones a realizar por los alumnos, se optó por el uso de un Animator Controller de Unity, aquí se incluirían todas las posibles animaciones a realizar por todos los alumnos. De esta manera más tarde en código se puede elegir qué animación ejecutar en cada momento dependiendo de la situación donde nos encontremos en la simulación.

Además, este método es fácilmente escalable y permite mantener las animaciones en loop cuando se requiere o evitar que pase de una animación a otra de manera inesperada.



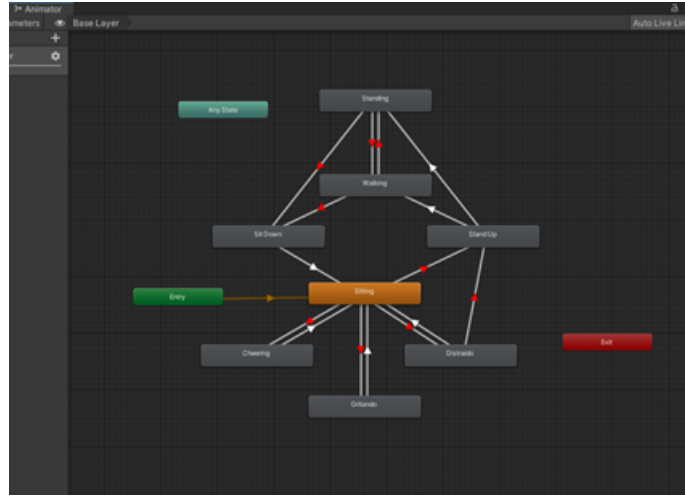


Figura 4.7: Esquema en Unity de las animaciones de los alumnos

### 4.2.3. Incorporación de Emopose

Basándonos en el proyecto [Emopose](#) [García-Magariño et al., 2019] hicimos programación inversa para entender y ser capaces de utilizarlo en nuestra herramienta. Una vez conseguido, cogimos las clases y componentes del proyecto que necesitábamos y creamos una escena nueva en nuestro proyecto. Esta escena fue la que utilizamos para entrenar el sistema con nuevas poses asociadas a emociones.

Las posiciones se guardaban tras el entrenamiento en un directorio denominado “learned poses” donde diferentes .txt, con el nombre de la emoción que representaban, guardaban una pose que asociaban a esa emoción. Estas posiciones no guardan la posición exacta de los miembros del usuario, sino que realizan una ponderación de las posiciones a través de [Inverse Kinematics] para poder guardar también los ángulos y posición exacta de los componentes de cada miembro.

La estructura utilizada para guardar las posiciones asociadas a una emoción está dividida en 8 valores diferentes separados por ”;” cada valor representa una parte del cuerpo en el siguiente orden: mano izquierda, mano derecha, pie izquierdo, pie derecho, cabeza, dirección de la mirada, apertura de la mano derecha y apertura de la mano izquierda.

En posible ejemplo del resultado al completar esta estructura es:

$(0.1662273, 1.084165, -0.1571718); (-0.121711, 1.033779, -0.1293866);$   
 $(0.05229819, -0.0156253, 0.193169); (-0.1103238, -0.01918733, -0.1278396);$   
 $(-0.000647052, 1.234938, 0.06255563); (-0.02433724, 0.9978769, -0.06041036); 1; 1$

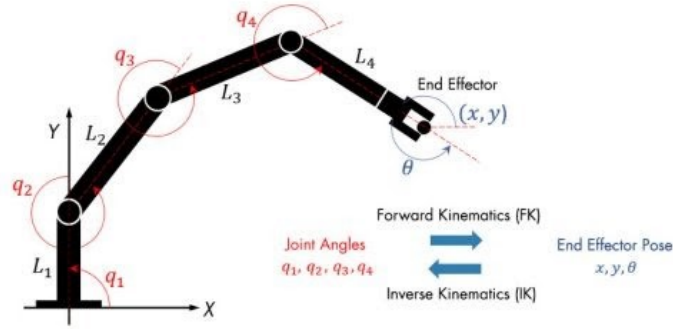


Figura 4.8: Funcionamiento de *Inverse Kinematics*

Para poder utilizar este proyecto en su máximo potencial hemos modificado la escena que reestructuramos en nuestro proyecto. De esta manera, el personaje ahora obtiene el input de [Axis Neuron](#) para colocarse en la posición que se desee grabar asociándola a una emoción y así facilitar la precisión necesaria en algunas poses.

#### 4.2.3.1. Problemas durante la incorporación

Inicialmente el personaje de la escena *Capture Pose* no tenía implementada la diferencia de distancia desde el punto inicial, para el cálculo de la emoción. Teniendo en cuenta el movimiento lateral que se produce al usar un traje, en lugar de mover únicamente los brazos, piernas y cabeza del personaje. Por lo que las posiciones recogidas tenían movimientos de posición entre -5 y 5, o incluso más altas, cuando deberían estar entre -2 y 2 como mucho.

Lo mismo pasaba con la escena de juego, no se tenía en cuenta las posiciones locales de cada parte del cuerpo del jugador, sino las globales. Esto era más llamativo cuando el usuario se movía por la clase, que es un espacio bastante amplio, provocando que los resultados de las emociones obtenidas estuvieran muy distorsionados en comparación a los entrenados previamente y, por tanto, las comparación frente a las emociones era muy amplia y poco fiable.

Para tratar de solucionar este problema se optó por mantener el avatar que representaba el cuerpo del jugador aislado del resto de la escena en las coordenadas (0,0). Esto nos permite minimizar el movimiento del avatar, ya que si te movías con los joysticks por la escena, el avatar no reproducía ese movimiento lateral. Sin embargo cuando andabas con el traje y las gafas de manera natural. Si que afectaba al avatar, y este se desplaza contigo.

El siguiente paso fue eliminar ese movimiento, para ello se realiza una ponderación entre la posición inicial, al comenzar la escena, y la actual, en el momento del análisis, para ignorar el movimiento lateral provocado por el andar del usuario.

Era necesario utilizar el sistema de reconocimiento de emociones a lo largo de la simulación de la situación. Por este motivo, también hubo que incorporar algunos elementos de EmoPose, como poseBuilder, a la escena *Class.GameScene*, además se modificó la lógica del creador de poses para que no tuviera en cuenta la posición global del jugador en el entorno de juego, sino la posición local de cada una de las extremidades. Más adelante se

encontró una modificación más acertada, basada en colocar el avatar, al cual se le capturaba la pose, alejado de la escena de juego. Esto permite evitar fallos en el avatar al moverte con los mandos, aunque la solución anterior seguía siendo necesaria ya que al moverte con el traje y las gafas, sin usar los joysticks de los mandos, el avatar también se mueve contigo.

También se creó la clase *Motion Capture Manager* que se encuentra en la escena *Clas\_GameScene* con la API necesaria para capturar las poses durante la ejecución del juego e ir guardando los resultados obtenidos en el archivo .csv correspondiente. Esta clase es principalmente se utilizan los métodos *init()* para inicializar los parámetros que necesitará la clase durante la ejecución y *update()* que cada determinado tiempo capturara la pose del usuario y la guardara, tanto internamente como en el *csv*.

#### 4.2.4. Adición de nuevas emociones

Una vez incorporada la herramienta [Emopose](#) [García-Magariño et al., 2019] en nuestra aplicación, procedimos a crear un pequeño Data Set con nuevas posiciones que estuvieran relacionadas con las emociones. Para ello realizamos una investigación sobre qué emociones podían verse reflejadas a través del lenguaje corporal.

##### 4.2.4.1. Investigación sobre el lenguaje corporal y las emociones

Como hemos comentado anteriormente en el capítulo del [Estado del Arte](#), la definición de comunicación no verbal varía según el autor al que nos refiramos. Sin embargo, la expresión corporal es un concepto claramente definido y sabemos que abarca los movimientos del cuerpo y postura del mismo. Siendo más concretos, según [Lhommet & Marsella, 2014], existen dos tipos de movimientos y pueden ser clasificados: movimientos faciales y movimientos gesticulares-posturales.

Una vez conocimos esta información analizamos por separado las dos categorías, obteniendo resultados muy variados que desarrollamos a continuación.

#### 1. Reconocimiento de emociones mediante expresiones faciales:

El sistema más popular para estimación de emociones mediante la expresión facial es “El método de codificación de Movimiento facial”, que fue descrito por Ekman y Friesen en 1978. Estos estudios se pueden consultar en libros como [Cohn et al., 2007].

Se trata de un sistema que, basándose principalmente en el movimiento de los músculos de la cara, puede estimar la emoción que el usuario está experimentando. Según los autores, el rostro humano puede ser descompuesto en unidades de acción (AUS) y cada una de ellas hace referencia a un músculo facial o a un conjunto de ellos. Mediante la combinación de diferentes AUS es posible estimar que emoción, de la [seis emociones básicas según Ekman](#), está experimentando el individuo.

Aunque la estimación de emociones por este método es bastante subjetiva si la realiza una persona, otros estudios como [Tarnowski et al., 2017] hacen ver que si el análisis lo realiza una computadora, resulta ser mucho más fiable.

Trabajos como *aproximación al reconocimiento de emociones faciales basado en posiciones de puntos de interés* de Víctor Manuel Álvarez y Ramiro Velázquez de 2018

[Pato & Guerrero, 2018], comienzan a hacer aproximaciones entre el campo de la psicología y la informática. Mediante este estudio establecieron los puntos de interés en rostros humanos que expresan emociones. Obteniendo como resultado los siguientes datos:

- **Alegría:** levantamiento de párpados inferiores y levantamiento de comisuras de los labios.
- **Tristeza:** elevación de la zona interior de las cejas y relajación de la comisura de los labios.
- **Sorpresa:** elevación excesiva de las cejas y de los músculos del párpado superior. Posible apertura de la boca y tensión en los labios.
- **Miedo:** según el autor, la más difícil de identificar, por la cantidad de músculos que participan en la expresión.
- **Ira:** el arco de las cejas decae y los párpados superiores caen.
- **Disgusto:** el punto central de la nariz sube ligeramente el labio parece tensarse con la forma de la boca completamente recta.

En busca de otros estudios que pudieran contrastar esta información, recurrimos a [Iglesias et al., 1989] y a [Sistema de codificación facial] donde además de confirmar estos datos, aportaron nuevos. De esta manera concretamos que las unidades de acción más representativas de la expresión facial son:

- **Alegría:** cejas levantadas, boca abierta enseñando los dientes, ascenso de las mejillas (AU6), elevación simétrica de la comisura de los labios (AU12) y arrugas en los laterales de los ojos.
- **Tristeza:** elevación de las cejas hacia el interior (AU1), caída de las comisuras de los labios (AU15), ascenso del mentón (AU17), párpados superiores caídos o mirada baja (AU41) y arrugas en el entrecejo.
- **Sorpresa:** elevación simétrica de las cejas hacia el exterior (AU2) e interior (AU1), apertura extrema de los párpados (AU5), caída de la mandíbula (AU26).
- **Miedo:** cejas bajadas (AU4), párpados en tensión (AU7), mirada hacia abajo (AU41), temblor en los labios, boca entreabierta (AU26), contracción del rostro hacia atrás (AU13), incremento de pestañeo por segundo (AU45) y retracción de los labios (AU20).
- **Ira:** cejas inclinadas hacia abajo (AU4), mirada fija, ojos y párpados tensos (AU41) y labios (AU20) y dientes apretados (AU27).
- **Disgusto:** arrugas en la nariz (AU09), elevación del labio superior (AU10), ceño fruncido, ojos entrecerrados o cerrados (AU43), mejillas ligeramente elevadas (AU6) y boca abierta (AU26), seguramente enseñando la lengua (AU19).

A continuación se muestran dos ejemplos de combinaciones de unidades de acción y la correspondiente emoción expresada.



Figura 4.9: Unidades de acción: AU1, AU2, AU5, AU26.  
Emoción expresada: sorpresa.



Figura 4.10: Unidades de acción: AU4, AU9, AU10.  
Emoción expresada: disgusto

## 2. Reconocimiento de emociones mediante expresiones corporales:

Aunque los estudios de Ekman sobre las [seis emociones básicas](#) indican que la cara es el canal más eficaz para expresar emociones, los gestos y posturas también dan información sobre el estado emocional como la excitación y tensión, datos que nos pueden ser identificados mediante la expresión facial. Estudios como [Wallbott, 1998] y [Lhommet & Marsella, 2014] han recopilado posturas que ocurren con frecuencia dependiendo de los estados emocionales.

- **Alegría:** extremidades ligeramente abiertas, postura erguida pero relajada, peso sobre una de las piernas, cabeza posiblemente ladeada, manos y brazo por encima de la cintura, palmas de las manos abiertas y dedos juntos.
- **Tristeza:** cabeza agachada, cuerpo echado ligeramente para atrás de tal manera que la cabeza sobresale, columna en forma de concha, cuerpo flácido, manos entrelazadas a la altura de la cadera, manos y brazo totalmente pegados al cuerpo.
- **Sorpresa:** manos situadas a los lados del cuerpo y por encima de la cintura con las palmas abiertas y dedos separados, peso repartido entre las dos piernas, es posible que se sitúen las manos en la cara.
- **Miedo:** cabeza agachada a la altura de los hombros, cruzar un solo brazo por delante para sujetar el otro, manos entrelazadas a la altura de la cintura, piernas cruzadas y tobillos juntos, por lo general, todo el peso del cuerpo se apoya en una sola pierna, rodillas ligeramente dobladas y cuerpo encorvado.

- **Ira:** puños cerrados, cuerpo echado para delante, manos agarradas a la espalda, brazos cruzados a la altura del pecho, cabeza levantada y la barbilla proyectada hacia delante, manos en los bolsillos con pulgares por fuera, manos en las caderas o separadas del tronco, piernas separadas y tensión general en el cuerpo con rodillas estiradas.
- **Disgusto:** hombros ligeramente echados atrás, cabeza mirando hacia abajo, cuerpo en torsión para evitar mirar al foco que produce esta emoción.

Los gestos faciales son los que mejor reflejan las emociones del emisor y por tanto, un análisis de las características faciales del usuario aportaría mejor resultado. Sin embargo, nuestro proyecto es un juego en Realidad Virtual. Esto implica que el usuario debe llevar unas gafas apropiadas para el uso de estas plataformas. En la actualidad, todas las gafas disponibles en el mercado tapan al menos un tercio de la cara, por lo que hace muy difícil hacer un análisis completo para obtener resultados óptimos.

Una vez descartado el seguimiento de los movimientos faciales, nos centramos en los movimientos gesticulares y posturales. De las conclusiones explicadas anteriormente, extrajimos diferentes combinaciones de todos los gestos y posturas pueden caracterizar una emoción. Además intentamos que estas posiciones estuvieran lo más diferenciadas posibles, para que el sistema de reconocimiento de [Emopose](#) no tuviera problemas a la hora de identificar la emoción. Algunos ejemplos de las combinaciones se muestran a continuación.



Figura 4.11: Combinaciones de expresiones corporales asociables una emoción. Alegría, disgusto, ira, miedo, sorpresa, tristeza y neutral respectivamente

#### 4.2.4.2. Preparación del equipo

El primer paso para conseguir recoger una emoción fue vincular el traje Perception Neuron, desde la plataforma Axis Neuron, con el ordenador. Para ello hay que establecer la conexión entre el traje y la aplicación, de manera que se permita el pase del streaming de datos de un lado a otro, existen dos formas: por cable de datos o por Wifi.

En nuestro caso, inicialmente elegimos la conexión por cable de datos por su facilidad de configuración y porque no precisamos alejarnos del ordenador para capturar movimientos. Esta conexión entre dispositivos se llevó como se explica en [Axis Neuron](#).

Tras establecer la conexión debemos calibrar el traje, para que ajuste sus neuronas al avatar proyectado en la aplicación. Una vez esté listo, tenemos preparado el entorno inicial para llevar a cabo la captura de posiciones.

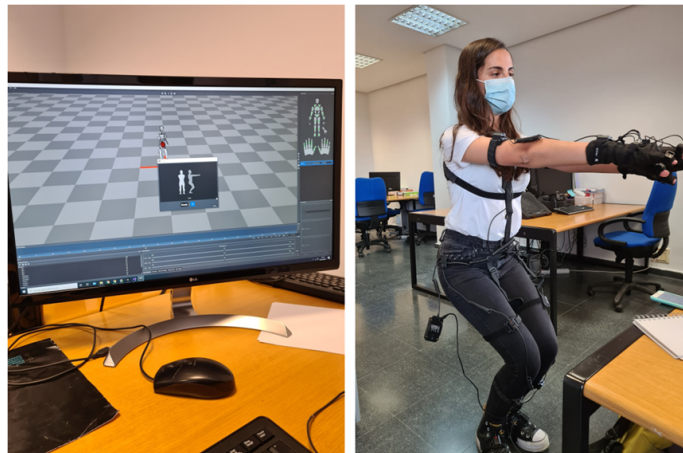


Figura 4.12: Calibrado del traje [Perception Neuron](#) mediante [Axis Neuron](#).

Una vez listo el sistema en Axis, procedimos a vincularlo a [Unity](#), tal y como explicamos en [Axis Neuron](#). Con el avatar de Unity recibiendo el streaming de datos proporcionado por Axis y, por tanto, moviéndose en la escena de juego imitando los movimientos del usuario, e integrada la escena con el proyecto de [Emopose](#). Se procedió a grabar las primeras poses asociadas a una emoción determinada.

Durante el proceso encontramos algunos problemas con [Emopose](#). Esta herramienta no estaba inicialmente diseñada para capturar poses de un avatar con movimientos laterales. De esta manera, solo podía capturar correctamente las emociones si el avatar se encontraba en la misma posición que cuando se grabó la posición.

Por este motivo, se tuvo que modificar el sistema para permitirnos ignorar la posición global de las extremidades y que solo tuviera en cuenta las locales. Aun así, no se ha conseguido una fiabilidad perfecta con la herramienta, ya que las gafas de VR también modifican la posición del avatar.



#### 4.2.4.3. Creación del Data Set de emociones

Inicialmente la base de casos de Emopose contaba con 6 emociones con 3 poses asociadas a cada una de ellas. Además, contaba con una emoción neutral, que era la considerada como ausencia de emoción.



Figura 4.13: Emociones asociadas a la expresión corporal [Schindler et al., 2008]

Tras el [estudio de las emociones](#) y la obtención de aquellas que creíamos obtendrían un mejor resultado (consultar imagen 4.11), procedimos a grabarlas e introducirlas en el Data Set con la emoción correspondiente.

En este sentido, hicimos diversas pruebas para comprobar con cuantas posiciones corporales por cada emoción obteníamos mayor efectividad de la herramienta. Primeramente, probamos con una única posición y a pesar de que todas ellas eran muy distintas, no lograba acertar por el escaso número de casos base. A continuación, probamos con cinco posiciones por emoción. En esta ocasión observamos que era difícil crear cinco posiciones distintas unas de otras por lo que la herramienta las confundía en la mayoría de ejecuciones.

Por último, probamos añadiendo tres posiciones por cada emoción, de esta manera el algoritmo tendría suficientes casos para comparar pero no se confundiría entre ellos porque serían lo suficientemente distintos. Concretamente, escogimos las tres primeras filas de la imagen 4.11. Para mejorar el resultado, añadimos dos casos más a la emoción neutral (en total cinco). De esta manera, pretendemos obtener emociones únicamente cuando la posición asociada a ella es muy relevante, en otro caso, sería neutral.



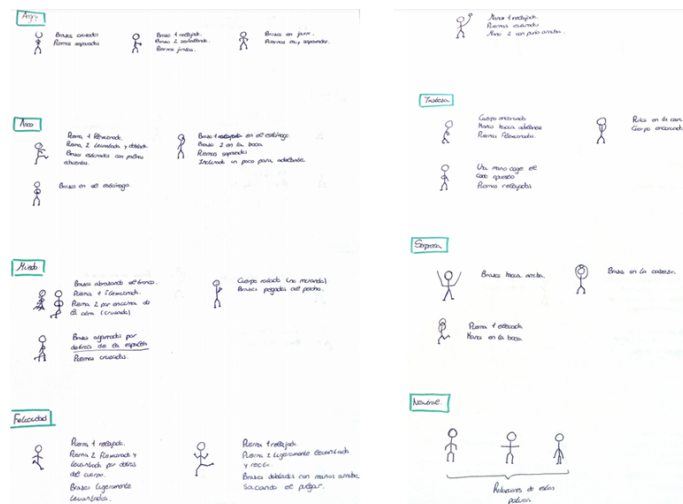


Figura 4.14: Esquema final de las emociones que conforman el Data Set de Classroom VR-Motion Capture

## Capítulo 5

# Experimentación

Una vez desarrollada la aplicación, comenzó la fase de pruebas. El día 9 de mayo se desplazaron a Barcelona dos integrantes del equipo de desarrollo junto con los tutores que coordinan el proyecto Didascalia además del Trabajo de Fin de Grado que se explica en este documento. Las pruebas comenzaron el día 10 de mayo a las 9:00h y concluyeron ese mismo día a las 18:30h.

### 5.1. Preparación

El lugar escogido para el desarrollo del experimento fue la Sala Polivalente de la Facultad de Educación de la Universidad Autónoma de Barcelona. Este aula es un espacio diáfano con sillas movibles. Esto nos ofreció un lugar espacioso para hacer uso de la VR además de cómodo para hacer los test y las entrevistas gracias a las numerosas sillas.



Figura 5.1: Sala Polivalente de la Facultad de Educación de la Universidad Autónoma de Barcelona

Primeramente, nos reunimos con los psicólogos, pedagogos y docentes del equipo Didascalia. Tras una breve introducción sobre el estado de la aplicación se propusieron una serie de mejoras rápidas que pudieron ser aplicadas antes de la llegada de los participantes del experimento.

Una vez encendido el equipo y conectados los dispositivos de captura de movimiento y realidad virtual, hicimos una pequeña prueba para comprobar que todos los cambios

incluidos funcionaban y estaba listo para el comienzo del experimento.

Por último, repartimos los roles que tomaría cada uno de los integrantes del equipo. De esta manera, cada uno sabríamos nuestro papel dentro del experimento y nos aseguraríamos de que el participante siempre estaría atendido por alguno de nosotros.

## 5.2. Participantes

Fueron citadas 14 personas para participar en el experimento. Todos ellos cuentan con experiencia en el campo de la educación. Entre ellos encontramos directores, profesores u orientadores de institutos de educación secundaria, profesores del Máster de Formación para el profesorado de Primaria o Secundaria y estudiantes del Máster de profesorado.

En cuanto a los años de experiencia como docentes, la amplia mayoría (57.1 %) llevan más de cinco años en el sector de la educación. Tan solo dos (14,3 %) de ellos cuentan con menos de dos años de experiencia.

Hablando de experiencias pasadas con la realidad virtual, el 64,3 % ya estaba familiarizado con el entorno, habiéndola utilizado en otras ocasiones. Los participantes restantes afirmaron que nunca habían probado un entorno de realidad virtual.



Figura 5.2: Gráficos extraídos del [cuestionario post-experimento](#) sobre datos demográficos de los participantes

## 5.3. Diseño experimental

Para el experimento, fueron reclutados y citados a lo largo del día 10 de mayo de 2021 los 14 candidatos mencionados anteriormente con un margen de 30 minutos. El experimento tiene una duración de 25 minutos y está compuesto de 4 etapas:

### 5.3.1. Explicación de la prueba

Los usuarios son recibidos por un miembro del personal técnico. Este le da la bienvenida y una vez sentados, le ofrece el [consentimiento informado](#) (incluido en el [anexo](#) de este documento) que, una vez firmado, permite al equipo la grabación de audio y vídeo con fines educativos.

A continuación, se proporciona al candidato un ID, aleatorio del 1 al 14, que será el que represente su identidad a lo largo del proceso. Con este procedimiento aseguramos el anonimato de los resultados de los participantes. Posteriormente, se procede a hacer unas preguntas de control como cuál es su experiencia profesional o si han utilizado en alguna

ocasión un dispositivo de realidad virtual. En caso de obtener una respuesta negativa a esta última pregunta, el técnico encargado de esta etapa, hace una pequeña explicación de los conceptos básicos de la realidad virtual. Haciendo énfasis en los [dispositivos](#) que se utilizarán y las principales sensaciones que pueden producir en el usuario (mareos, dolores de cabeza o posibles caídas).

Una vez recogida esta información, se hace saber al usuario que puede detener la prueba y abandonar el experimento en cualquier momento con solo comunicárselo a la persona que le esté asistiendo en ese momento. A continuación, el técnico da una breve explicación de las etapas que conforman el experimento y el recorrido que deberá seguir por la sala hasta completarlo.

Por último, se pregunta si tiene alguna duda sobre el procedimiento. En caso negativo, se le acompaña a la siguiente etapa: el desarrollo de la prueba.

### 5.3.2. Desarrollo de la prueba

Una vez llega el usuario a la estación donde se encuentra todo el [material](#) para hacer la inmersión de realidad virtual, se procede a su equipación con todos ellos.

Primeramente, se coloca el [Perception Neuron](#). El dispositivo cubre las extremidades superiores e inferiores, incluyendo las manos y los pies además del tronco. Una vez que el usuario se siente cómodo con el traje, lo calibramos. Este proceso es necesario para sincronizar cada parte del cuerpo con el avatar que utilizará el simulador durante la ejecución. A continuación, se colocan las [Oculus Rift S](#) y se ajustan a la cabeza del usuario. Por último, se coloca un auricular en uno de sus oídos para que pueda escuchar lo que se dice dentro del juego y, al mismo tiempo, las instrucciones del personal técnico de manera clara por si fuera necesario.

Una vez comprobada la conexión de los dispositivos con la aplicación, comienza la prueba. Para el análisis posterior que realizará el equipo técnico, comienza una grabación de la prueba con una cámara externa. A continuación, se ejecutará la aplicación principal y se pide al usuario que se familiarice con el entorno caminando y mirando a su alrededor. Esta práctica facilita y mejora la posterior experiencia de usuario. Una vez que se siente cómodo, el técnico encargado de esta etapa, le pide que observe el Menú Inicial y pulse el botón “Jugar”.

Tras pulsar el botón citado, se muestra el Menú de Escenas. Este muestra las tres situaciones a las que se enfrentará el usuario. Durante la prueba, se llevarán a cabo las tres por orden.

Cuando entra en una de las situaciones se muestra una breve descripción del entorno en que se encuentra: “entras en el aula, los alumnos están en sus grupos de trabajo correspondientes, puedes continuar con la explicación”. Una vez que el usuario se ha situado hay un tiempo en el que se espera que simule dar clase como si estuviese en un aula real. Pasado ese tiempo se produce una situación disruptiva: “un alumno dice que no quiere seguir trabajando con su grupo”, “un alumno se levanta y grita”, etc. Acto seguido se muestra una lista con las posibles opciones que el usuario puede hacer para solventar el conflicto: “dile que se siente y no moleste”, “acércate a él y cálmale”, etc. El jugador debe llevar

a cabo una de esas acciones de manera natural. Para ello, el participante podrá moverse libremente por el escenario, gesticular, hablar en voz alta, etc. El juego reconocerá uno de los caminos tomados analizando las acciones que ha llevado a cabo el participante y mostrará un feedback final.

Esta retroalimentación consiste en varias pantallas y alude a distintos ámbitos de la actuación:

- La primera da una breve explicación de si la acción elegida para solventar la situación conflictiva era la más adecuada o no. En caso de no ser la más efectiva, se hace un pequeño análisis de lo que hubiese pasado el aula a continuación de realizar la acción.
- La segunda pantalla muestra la emoción predominante asociada a las posturas captadas mediante el traje durante la escena. Esta puede ser una de las seis emociones básicas: miedo, ira, alegría, asco, sorpresa o tristeza. En caso de no reconocerse ninguna de las últimas, la emoción será neutral.
- Por último, la tercera pantalla refleja la diferencia en la tonalidad que ha utilizado el jugador durante la interacción con los alumnos virtuales. Esa se muestra mediante un porcentaje.

Este proceso se repetirá con las tres situaciones principales que se muestran en el Menú de Escenas. Una vez se acabadas, se retiran las [Oculus Rift S](#) y el traje [Perception Neuron](#) al usuario y se detiene la grabación de la prueba. Mientras se desinfectan los dispositivos para futuros usuarios, se acompaña al participante a la siguiente estación: el cuestionario post-prueba.

### 5.3.3. Cuestionario post-prueba

En esta etapa del experimento se ofrece al usuario un ordenador con un cuestionario basado en el Modelo de Aceptación de Tecnología (TAM, [Davis, 1989]). Estos cuestionarios han sido ampliamente utilizados en investigación educativa. En nuestro caso, adaptamos el cuestionario propuesto por [Hwang et al.2013], en un estudio similar al nuestro, en Japón. Se incluyeron un total de 18 ítems agrupados en las cuatro dimensiones que define el TAM:

- **Utilidad percibida:** se define como el grado en que el usuario (docente/docente en formación) percibe que el uso de la herramienta puede fomentar su competencia para gestionar conflictos de aula (ocho ítems). Por ejemplo: “Percibí que mi comportamiento - movimientos, actitudes, palabras... impactó a los estudiantes en el entorno virtual.”
- **Facilidad de uso percibida:** es el grado en que un usuario (docente/docente en formación) percibe que el uso de la herramienta no le supondrá ningún esfuerzo” (cuatro ítems). Por ejemplo: “Los dispositivos (gafas y traje) no han interferido en las acciones que deseaba realizar.”
- **Actitud hacia el uso:** se refiere a las emociones (positivas o negativas) experimentadas por el usuario (docente/docente en formación) en su experiencia en con

la herramienta (tres ítems). Por ejemplo: “Es divertido practicar la competencia de gestión de clima del aula en este entorno virtual.”

- **Intención conductual de usar la herramienta:** o probabilidad de adoptar la herramienta como un entorno para enseñar y aprender la competencia de gestión de clima de aula (tres ítems). Por ejemplo: “Mi competencia docente para gestionar el clima de aula puede mejorar con el uso continuo de este entorno de aprendizaje virtual.”

Para expresar la respuesta se utiliza una escala tipo likert de siete niveles, con la siguiente correspondencia: (1) totalmente en desacuerdo; (2) parcialmente en desacuerdo (3) en desacuerdo; (4) neutral, (5) parcialmente de acuerdo; (6) de acuerdo; (7) totalmente de acuerdo. El cuestionario se administra en línea, al final de la prueba, a través de un formulario de Google. El cuestionario se incluye en el apartado [Preguntas del cuestionario post-test](#) del [Anexo](#)

#### 5.3.4. Entrevista final

La última etapa consiste en una entrevista semi estructurada con un componente del equipo investigador de aproximadamente 10 minutos. Esta es recogida mediante grabadoras de audio para su posterior análisis y extracción de resultados. En ella se explora la valoración de la experiencia por parte de los participantes, específicamente sobre el feedback proporcionado en cuanto a emociones que sintió durante la ejecución de la prueba del sistema. Además, se solicitaron sugerencias sobre posibles mejoras de la herramienta, las cuales son claves para el trabajo futuro que presentamos más adelante en el capítulo 11.

A continuación, se muestra una imagen del procedimiento que siguieron los usuarios para completar el experimento.

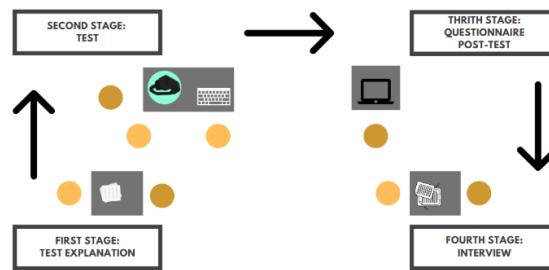


Figura 5.3: Esquema de las etapas del experimento

#### 5.4. Correcciones

Tras concluir las pruebas, nos dimos cuenta de ciertos elementos que requerían cambios, bien por fallos o bien por posibles mejoras.

La lista que se muestra a continuación son todas correcciones o cambios realizados después de la fase de experimentación realizada en Barcelona:

- **Eliminar los paneles iniciales y finales de la escena del juego, y sustituirlo por audios:**

Se han grabado audios nuevos para cada uno de los contextos aportados por las escenas, así como para cada uno de los posibles feedbacks finales que se dan al final de las escenas dependiendo del camino elegido. Además, se ha añadido el script `StartPlaying.cs`, encargado de seleccionar el componente `AudioSource` correspondiente de cada jugador, dependiendo de la plataforma en la cual se esté ejecutando la aplicación.

- **La mirada de los alumnos seguirá al profesor por donde se mueva, aumentando así el realismo de la aplicación.**

Para realizar esto, se añadió un componente `Rig Builder` a cada prefab de alumnos, un objeto `HeadRig`, con un componente `Rig` que debe asignarse al `Rig Builder` mencionado, y un objeto `HeadAim` con un componente `MuiltiAim Constraint` que debe tener asignado el objeto head del alumno como `constrainedObject` y el `GameObject` del jugador como `sourceObject`. Para lograr esto, se ha añadido el script `SelectAim-Target`, que selecciona como `sourceObject` al jugador correspondiente, estos scripts diferencian la plataforma en la que se ejecuta la aplicación.

- **Mejorar el comportamiento de los botones en los menús para facilitar su pulsación.**

Se ha aumentado el tamaño del `RectTransform` de los botones, para aumentar su área de acción sin afectar al tamaño del texto.

- **Mejorar la detección de audio**

Se ha incluido el script `SoundManager` que contiene las referencias a los scripts que manejan el audio, que son `MicToVokaturi`, `MicrophoneManager`, `KeyWordRecognizer` y `SoundLoudness`. Este script inicializa `KeyWorkdRecognizer`, y procesa la información de `Vokaturi` para la serialización mediante una llamada al método `processVokaturiInfo()`. Las llamadas de `MySceneManager` a cualquiera de estos componentes se realizan mediante la instancia de `SoundManager` que guarda.

- **En la segunda escena (Unión problemática) de Classroom VR debe comenzar directamente en el estado de detección de camino.**

Se ha editado el flujo de la escena en `MySceneManager` para detectar situaciones con un valor de `timeToStart` de -1 para diferenciar entre escenas que tienen o no presentación de la situación, permitiendo así mostrar escenas cuya intervención del usuario es requerida desde el inicio.

## Capítulo 6

# Resultados

Tras la [fase de experimentación](#) llevada a cabo en Barcelona, hicimos un estudio de los resultados que obtuvimos con las respuestas de los usuarios. Estos resultados los dividimos y analizamos en tres partes.

Primeramente se muestran los [resultados obtenidos por CVR-MC](#), su estructura y un ejemplo de los mismos. A continuación, se reflejan los [resultados cuantitativos](#). La información para crear este estudio de los resultados son los recogidos por el [cuestionario](#) que rellenaron los usuarios tras finalizar la prueba. Por lo tanto, los resultados de este apartado reflejan la usabilidad percibida por los usuarios. Seguidamente se muestran [resultados cualitativos](#). Esta información es la recogida de las respuestas de los usuarios a la [entrevista final](#). La clasificación de estos resultados fue realizada con ayuda de Ibis Álvarez. Este proceso consistió en la transcripción de todas y cada una de las respuestas y su posterior clasificación y agrupamiento para obtener los resultados cualitativos finales.

### 6.1. Resultados recogidos por la aplicación CVR-MC

Gracias al sistema de serialización implementado, explicado en el apartado de [Estructura y generación de situaciones](#), fuimos capaces de recoger diversos datos durante la prueba piloto realizada. Estos se guardan en formato .csv para visualizarlos más fácilmente, poder analizarlos en el futuro y utilizarlos en el entrenamiento de futuros sistemas con sistemas de inteligencia artificial más complejos (que suelen necesitar un gran número de datos). Entre estos datos podemos encontrar los siguientes:

- Emoción detectada.
- Posición mano izquierda.
- Posición mano derecha.
- Posición pierna izquierda.
- Posición pierna derecha.
- Posición cabeza.
- Posición dirección de vista.



- Apertura de la mano izquierda.
- Apertura de la mano derecha.
- Diferencia entre la posición del Data Set y la posición asociada a la emoción detectada.

Además, al final se añaden unos pequeños textos indicando las palabras clave reconocidas por el sistema mediante la detección de voz, las variaciones en la tonalidad durante el discurso y el tiempo total empleado en resolver la situación.

Consideramos que esta información recogida sumada a las grabaciones realizadas durante el desarrollo de la prueba pueden sentar la base de futuras investigaciones en este campo. A continuación se muestran algunos ejemplos de los datos recogidos en el formato procesado.

Emocion	Mano Izquierda	Mano Derecha	Pierna Izquierda	Pierna Derecha	Cabeza	Dirección de Visión	Apertura mano	Apertura mano	Distancia	
Happy	(-0.4/ 1.0/ 0.0)	(-0.7/ 1.2/ -0.2)	(-0.7/ 0.1/ 0.0)	(-0.5/ 0.1/ -0.1)	(-0.5/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	41.4	
Happy	(-0.4/ 1.1/ 0.0)	(-0.8/ 1.2/ -0.2)	(-0.7/ 0.1/ 0.0)	(-0.5/ 0.1/ -0.1)	(-0.5/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	41.86	
Happy	(-0.4/ 1.1/ 0.0)	(-0.9/ 1.1/ -0.1)	(-0.7/ 0.1/ 0.0)	(-0.5/ 0.1/ -0.1)	(-0.6/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	43.59	
Happy	(-0.4/ 1.1/ 0.0)	(-0.9/ 1.1/ -0.1)	(-0.7/ 0.1/ 0.0)	(-0.6/ 0.1/ -0.1)	(-0.6/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	44.05	
Happy	(-0.5/ 1.0/ 0.0)	(-0.9/ 1.1/ -0.1)	(-0.7/ 0.1/ 0.0)	(-0.6/ 0.1/ -0.1)	(-0.6/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	45.88	
Fearful	(-0.4/ 1.0/ 0.0)	(-0.9/ 1.1/ 0.0)	(-0.6/ 0.1/ 0.0)	(-0.6/ 0.1/ -0.1)	(-0.6/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	44.44	
Fearful	(-0.3/ 1.1/ 0.0)	(-0.9/ 1.1/ -0.1)	(-0.7/ 0.1/ 0.1)	(-0.6/ 0.1/ -0.1)	(-0.6/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	44.32	
Sad	(-0.4/ 1.1/ -0.2)	(-0.9/ 1.1/ 0.0)	(-0.7/ 0.1/ 0.1)	(-0.6/ 0.1/ -0.1)	(-0.6/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	45.34	
Sad	(-0.4/ 1.1/ -0.1)	(-0.9/ 1.1/ -0.1)	(-0.7/ 0.1/ 0.1)	(-0.6/ 0.1/ -0.1)	(-0.6/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	45.56	
Sad	(-0.3/ 1.1/ -0.1)	(-0.9/ 1.2/ -0.2)	(-0.7/ 0.1/ 0.1)	(-0.6/ 0.1/ -0.1)	(-0.6/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	45.07	
Sad	(-0.5/ 1.1/ -0.3)	(-0.9/ 1.1/ 0.0)	(-0.7/ 0.1/ 0.0)	(-0.6/ 0.1/ -0.1)	(-0.6/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	45.16	
Sad	(-0.7/ 1.1/ -0.3)	(-0.9/ 1.1/ 0.1)	(-0.6/ 0.1/ 0.0)	(-0.7/ 0.1/ -0.1)	(-0.7/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	49.27	
Neutral	(-0.8/ 1.1/ -0.2)	(-0.7/ 1.1/ 0.2)	(-0.6/ 0.1/ 0.0)	(-0.7/ 0.1/ -0.1)	(-0.7/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	46.52	
Neutral	(-0.8/ 1.1/ -0.3)	(-0.6/ 1.1/ 0.1)	(-0.6/ 0.1/ 0.0)	(-0.7/ 0.1/ -0.1)	(-0.6/ 1.5/ -0.2)	(0.0/ 0.0/ 0.0)	0.33	0.33	42.93	
Neutral	(-0.8/ 1.2/ -0.2)	(-0.7/ 1.1/ 0.1)	(-0.6/ 0.1/ 0.0)	(-0.7/ 0.1/ -0.1)	(-0.7/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	42.85	
Neutral	(-0.9/ 1.2/ -0.1)	(-0.7/ 1.0/ 0.2)	(-0.6/ 0.1/ 0.0)	(-0.7/ 0.1/ -0.1)	(-0.7/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	41.68	
Neutral	(-1.0/ 1.2/ -0.2)	(-0.7/ 1.1/ 0.2)	(-0.6/ 0.1/ 0.0)	(-0.7/ 0.1/ -0.1)	(-0.8/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	43.86	
Neutral	(-0.8/ 1.1/ -0.2)	(-0.9/ 1.2/ 0.2)	(-0.6/ 0.1/ 0.0)	(-0.7/ 0.1/ -0.1)	(-0.8/ 1.5/ 0.0)	(0.0/ 0.0/ 0.0)	0.33	0.33	47.56	
Neutral	(-0.9/ 1.2/ -0.2)	(-0.8/ 1.2/ 0.2)	(-0.6/ 0.1/ 0.0)	(-0.8/ 0.1/ -0.1)	(-0.8/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	47.87	
Neutral	(-0.9/ 1.2/ -0.2)	(-0.7/ 1.2/ 0.2)	(-0.6/ 0.1/ 0.0)	(-0.9/ 0.1/ -0.1)	(-0.8/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	46.32	
Neutral	(-0.9/ 1.1/ -0.3)	(-0.6/ 1.1/ 0.1)	(-0.6/ 0.1/ -0.1)	(-0.9/ 0.1/ -0.1)	(-0.8/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	46.38	
Sad	(-0.9/ 1.1/ -0.3)	(-0.7/ 1.1/ 0.1)	(-0.9/ 0.1/ -0.2)	(-0.9/ 0.1/ -0.1)	(-0.8/ 1.5/ -0.2)	(0.0/ 0.0/ 0.0)	0.33	0.33	47.75	
Sad	(-1.0/ 1.1/ -0.2)	(-0.8/ 1.1/ 0.2)	(-0.9/ 0.1/ -0.2)	(-0.9/ 0.1/ -0.1)	(-0.9/ 1.5/ -0.2)	(0.0/ 0.0/ 0.0)	0.33	0.33	53.23	
Sad	(-1.1/ 1.1/ -0.2)	(-0.9/ 1.1/ 0.2)	(-0.8/ 0.1/ -0.2)	(-0.9/ 0.1/ -0.1)	(-1.0/ 1.5/ -0.1)	(0.0/ 0.0/ 0.0)	0.33	0.33	56.85	

Figura 6.1: Primera parte del CSV resultado de la simulación del participante 2 (P02)

Emocion a	Y	Mano Izquierda	Mano Derecha	Pierna Izquierda	Pierna Derecha	Cabeza	Y	Direccion de/ve	Apertura mano	Apertura mano	Distancia Y
Sad		(0.6/ 1.1/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.66
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.3)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.3)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.94
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.3)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.3)		(0.0/ 0.0/ 0.0)	0.33	0.33	59.5
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.3)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.3)		(0.0/ 0.0/ 0.0)	0.33	0.33	59.08
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.3)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.67
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.3)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.64
Sad		(0.7/ 1.2/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.62
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.23
Sad		(0.7/ 1.1/- 1.4)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.59
Sad		(0.7/ 1.2/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.27
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.41
Sad		(0.7/ 1.2/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	59.1
Sad		(0.6/ 1.2/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.18
Sad		(0.6/ 1.2/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.17
Sad		(0.7/ 1.2/- 1.4)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.46
Sad		(0.7/ 1.2/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	57.94
Sad		(0.7/ 1.2/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	57.86
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	57.96
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.15
Sad		(0.7/ 1.1/- 1.4)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	59.03
Sad		(0.7/ 1.2/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.33
Sad		(0.7/ 1.1/- 1.4)	(0.8/ 0.9/- 1.3)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	59.44
Sad		(0.7/ 1.1/- 1.4)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	59.09
Sad		(0.7/ 1.2/- 1.5)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	59.91
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	57.95
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	57.83
Sad		(0.7/ 1.1/- 1.4)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	59.71
Sad		(0.7/ 1.1/- 1.3)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	57.93
Sad		(0.7/ 1.1/- 1.4)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.5/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.39
Sad		(0.6/ 1.1/- 1.4)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.85
Sad		(0.7/ 1.1/- 1.4)	(0.8/ 0.9/- 1.2)	(0.9/ 0.1/- 1.2)	(0.7/ 0.1/- 1.2)	(0.7/ 1.4/- 1.2)		(0.0/ 0.0/ 0.0)	0.33	0.33	58.61
Se ha reconocido. Despues											
Tiempo en resolver la situación: 64		39 segundos									

Figura 6.2: Segunda parte del CSV resultado de la simulación del participante 2 (P02)

## 6.2. Resultados cuantitativos de la prueba de usabilidad del sistema

Según el modelo usado para la creación del test [Davis, 1989], hay cuatro variables principales que afectan la aceptación y la adopción de nuevas tecnologías por parte de los usuarios: facilidad de uso percibida, la utilidad percibida, actitud hacia el uso e intención de uso. Para expresar la respuesta se utiliza una escala tipo likert de siete niveles, desde (1) totalmente en desacuerdo hasta (7) totalmente de acuerdo. A continuación, se incluyen las preguntas que conformaron el cuestionario post-prueba (basado en TAM) que rellenaron los usuarios tras la prueba con CVR-MC.

### 1. Utilidad percibida (UP)

- ¿Tienen los estudiantes del entorno apariencia y actitudes humanas?
- ¿Es natural la actitud de los estudiantes cercanos al conflicto?
- El número de estudiantes es apropiado para los ejercicios planteados.
- Los alumnos respondieron correctamente a mis acciones. Percibí que mi comportamiento provocó un impacto a los estudiantes.
- Los estudiantes entienden lo que decía.
- Se despertaron emociones en mi hacia los estudiantes. Pude adecuar mi forma de gestionar el conflicto a la situación presentada.

### 2. Facilidad de uso percibida (UF)

- Me sentí como si estuviera frente a una clase.
- Me resultó fácil manejar la herramienta.

- c) Me sentí cómodo utilizando los dispositivos de VR
- d) Las gafas y el traje no han interferido en las acciones que deseaba realizar.

### 3. Actitud hacia el uso (UA)

- a) He disfrutado el uso de la herramienta.
- b) Es divertido practicar la gestión del clima del aula con esta herramienta.
- c) La práctica en este entorno virtual me parece emocionante.

### 4. Intención de uso (IU)

- a) La simulación ayuda al aprendiz a mejorar su habilidad en el manejo de incidentes.
- b) Mi competencia docente puede mejorar con el uso continuo de esta herramienta.
- c) Sería genial contar con esta herramienta para apoyar la formación de los docentes.

Las preguntas de la a a la h nos muestran información sobre la utilidad de CRV-MC según los usuarios. Para su análisis, sumamos las respuestas obtenidas. Para poder sumar estas preguntas se realizó un alpha de cronbach, obteniendo como resultado un alpha de 0.815 y una media de respuestas de 4.7 sobre 7. Estas respuestas nos dicen que los usuarios consideran que la herramienta puede mejorar su competencia para gestionar conflictos en el aula. Esto coincide con la gran acogida que mostraron los usuarios en la entrevista final y con la necesidad de obtener un entorno seguro para la práctica de estas competencias.

Por otra parte, las preguntas i, j, k, l corresponden a la facilidad de uso percibida. Para analizar las respuestas a estas preguntas se vuelve a crear un alpha de cronbach con la suma de todas ellas. Como resultado obtenemos un alpha de 0.84 y una media de 5.1 sobre 7, mostrando que gran parte de los usuarios no les supondría esfuerzo utilizar la herramienta, sin embargo, se notaba cierta incomodidad. En las entrevistas pudimos conocer las limitaciones que detectaron los participantes.

A continuación analizamos las preguntas de la m a la o, que muestran información sobre la actitud hacia el uso. Para analizar los resultados nos ayudamos de un alpha de cronbach que dio como resultado 0.802 y una media de 6.1. Estos resultados muestran que la experiencia vivida con la herramienta ha sido relativamente buena. Con estos datos observamos que aunque los entornos de realidad virtual son algo novedoso en este sector, también son emocionantes y llamativos.

Por último, las preguntas p, q y r refieren la intención de uso de la herramienta por parte de los usuarios. Mediante un alpha de cronbach podemos analizar los resultados a estas preguntas. En este sentido hemos obtenido un alpha de 0.617 y una media de 6.3 sobre 7. Con esto y el apoyo de las entrevistas, podemos concluir que la herramienta con ciertas mejoras podría tomar papel como uso complementario en la formación del profesorado.

A continuación se muestra una tabla con el estudio de los cuatro campos principales mediante [Alpha de Cronbach]

Grupo de preguntas	Número de elementos	Alpha de Cronbach	Alpha basada en elementos normalizados
UP	8	0.815	0.813
UF	4	0.84	0.864
UA	2	0.802	0.841
IU	3	0.617	0.689

Cuadro 6.1: Resumen de los resultados obtenidos en el cuestionario post-prueba para valorar la adaptación de los usuarios a CVR-MC (escala tipo likert de siete niveles)

### 6.3. Resultados cualitativos de la entrevista

El análisis del contenido de la entrevista posterior a la experiencia se guió principalmente por las siguientes preguntas:

- ¿Qué aspectos valoran los participantes en su experiencia como docentes en formación en el CVR-MC?
- ¿Qué obstáculos percibieron en el diseño tecno-pedagógico del escenario formativo del CVR-MC?
- ¿Cuáles son las potencialidades percibidas por los participantes en la prueba del sistema CVR-MC para su utilización en la formación docente?
- ¿Cuáles son los aspectos que se deberían mejorar para adecuar el sistema a los requerimientos de una formación docente dirigida a fomentar la competencia para la gestión del clima de aula?

Se presenta a continuación una síntesis de los aspectos más relevantes, en relación con cada una de estas cuestiones.

En líneas generales, los aspectos más destacados fueron dos. Por un lado, la experiencia inmersiva en el rol docente frente a un grupo de estudiantes y, por otro lado, el realismo del escenario para practicar la competencia de gestión del clima del aula.

*“Me ha gustado la experiencia de creerte, de verte en un aula real. Aunque fuera realidad virtual, me ha despertado credibilidad. Enseguida les he dado a los alumnos virtuales una psicología: este no está escuchando, aquel está distraído, aquel me va a dar problemas (P02).”*

Además, algunos participantes destacaron el carácter lúdico de la experiencia. Respecto a los obstáculos, la mayoría de los participantes en la prueba manifestó desagrado con la presentación de opciones (camino) para modelar su actuación cuando ocurrían conductas disruptivas. En su opinión, de esta manera se limitan las posibilidades de aprendizaje.

*“El hecho de que se marque en color rojo que se va a producir el problema o con qué alumno se va a producir el problema, eso quizás te prepara un poco para ver que el problema va hacia allá y orienta un poco la respuesta (P03).”*

También hicieron referencia a limitaciones para la interactividad y dificultades iniciales para manejar la tecnología, echando en falta un entrenamiento previo para familiarizarse con el sistema y con los dispositivos.

No obstante, todos los participantes en la prueba consideraron que el sistema es útil para la formación docente, destacando, sobre todo, su orientación a la práctica.

*“Me parece más significativo que otras experiencias formativas que conozco en el sentido de que se puede estudiar directamente una situación, haciendo un estudio de caso y pensando qué harías. Me parece mucho más significativo cuando te pones, cuando lo experimentas, claro (P01).”*

Las potencialidades percibidas se refieren tanto a la relevancia del contenido como a la innovación metodológica. Muchos comentarios subrayan la necesidad de incluir la gestión del clima del aula en los programas formativos, sobre todo, para los profesores noveles.

Específicamente, respecto a la utilidad del sistema para el aprendizaje de la competencia de gestión de clima de aula, se destacó la posibilidad de trabajar los aspectos emocionales y actitudinales. Por ejemplo:

*“Para mí lo valioso de esta propuesta es que ellos [los docentes en formación] se pongan en contacto con sus propias emociones, con sus temores, con sus miedos. En las clases [del máster de secundaria] nos dicen [a los docentes] que les preocupa mucho enfrentarse al grupo, que les da miedo hacerlo mal, no estar a la altura (P09).”*

Se visualizó también la posibilidad de construir conocimiento de manera colaborativa, derivada del análisis posterior de la práctica en el entorno virtual.

*“Es una experiencia necesaria, no sólo en el momento de lo que tú has hecho y como tú te has sentido en el plano individual, sino también estaría bien comentarlo con los demás. Por ejemplo, tomar cada una de estas situaciones y a los que han pasado por allí [docentes en formación] y preguntar ¿Tú qué has dicho y por qué has hecho esto? ¿Y si hubieras hecho esto otro? ...Yo creo que la posibilidad de compartirlo con otros puede ser aún más interesante (P03).”*

## Capítulo 7

# Discusión

En esta sección responderemos a las dos preguntas de investigación planteadas. Para ello, utilizaremos los [resultados](#) obtenidos y explicados en el apartado anterior.

**RQ1:** *¿Es factible utilizar el sistema CVR-MC en la formación docente para contribuir al aprendizaje de competencias comunicativas de gestión de clima de aula?*

Sí, tras analizar los resultados podemos afirmar que nuestra herramienta tiene el potencial para funcionar como un sistema de aprendizaje, capaz de contribuir a la formación del profesorado. Los usuarios coinciden en que puede ser una herramienta de apoyo muy recomendable para los profesores en proceso de formación. CVR-MC les permitirá practicar, en un espacio seguro, las técnicas necesarias para afrontar una situación conflictiva, lo que permitirá un mayor desarrollo en sus habilidades de resolución de conflictos.

Acorde a los comentarios de los expertos, el feedback proporcionado por CVR-MC es interesante para el aprendizaje del usuario. De la misma manera, un debate conjunto durante la formación docente puede ser muy constructivo e interesante para los futuros profesores.

Sin embargo, también pudimos observar cómo algunos participantes hicieron referencia a problemas de aprendizaje de la herramienta por manejabilidad donde podemos deducir que un entrenamiento previo al uso práctico sería necesario.

**RQ2:** *¿Es posible capturar el lenguaje no verbal y las emociones que manifiestan los participantes durante la simulación? ¿Se corresponden con las que manifiestan en un entorno real?*

No. La asociación de emociones a las posiciones que adoptaron los usuarios no alcanzó el resultado esperado. El motivo principal es que los usuarios no se movían como se esperaba en un inicio y no se detectaban tantas emociones como se pretendía.

Si bien el sistema de captura de movimiento y reconocimiento de emociones funciona perfectamente en un entorno teórico, en la práctica hemos podido observar que, como la expresión corporal de los usuarios se veía comprometida, Emopose no pudo hacer un análisis óptimo de las emociones. En consecuencia, aunque sí es posible la captura del lenguaje no

verbal y estimar la emoción principal del usuario, no podemos hacer un seguimiento de estas posiciones si los usuarios no muestran su lenguaje corporal.

Tras reflexionar sobre el motivo del comportamiento de los usuarios, deducimos a partir de las entrevistas, que la incomodidad del equipamiento que llevaban los participantes (gafas, mandos, y traje de captura de movimiento) hizo que sus gestos fueran menos significativos que los que suelen hacer en la vida real. Esto ha hecho que las emociones detectadas por el sistema Emopose, (un sistema no basado en entornos virtuales) no hayan sido las que los participantes decían experimentar. Creemos que es necesario realizar un estudio más exhaustivo del movimiento corporal en entornos virtuales ya que parece ser diferente al utilizado en entornos reales.

## 7.1. Limitaciones del estudio

Una de las limitaciones encontrada durante el desarrollo de la herramienta fue el análisis de las posiciones generadas por el traje [Perception Neuron](#) y la estimación de la emoción que el docente transmitía durante la simulación. El principal motivo es que la herramienta EmoPose, encargada de la asignación de emociones, no estaba planteada inicialmente para usarse con una persona en movimiento sino con un avatar estático.

Por otra parte, antes de comenzar el desarrollo de CVR-MC, intentamos obtener un prototipo de [gafas de Realidad Virtual](#) de los laboratorios [Emteq labs](#). Estas, hubieran sido una fantástica opción ya que permiten obtener información relacionada con la dirección de la mirada y algunos parámetros relacionados con el funcionamiento del cerebro del usuario. Toda esta información habría sido, sin duda, muy valiosa para poder tener una perspectiva más amplia de lo que siente el docente durante el desarrollo de la situación. Desafortunadamente no se pudo obtener esta tecnología y mantuvimos el desarrollo programado con las [Oculus](#).

Asimismo, dado que nuestros conocimientos relacionados con el campo de lenguaje no verbal no son tan amplios como hubiera sido deseable y, teniendo en cuenta la limitación de tiempo y tamaño de grupo, no se han podido mejorar funcionalidades extras en la herramienta. Como el reconocimiento de voz y su tonalidad.

Nos gustaría aclarar que si todos estos módulos (reconocimiento de voz, captura de movimiento, dirección de la mirada y análisis de ondas cerebrales) se implementan en la versión final de CVR-MC y se utiliza toda esta información para la estimación de emociones, la aplicación tiene una amplia capacidad de usarse para contribuir en el aprendizaje de los docentes. No solo por sí misma, sino también como apoyo durante las sesiones de formación ya que puede promover el diálogo y la discusión sobre qué estrategia de gestión de conflicto es la más acertada.

Finalmente queremos concluir que sin duda el análisis del lenguaje corporal mediante nuevas tecnologías no está tan desarrollado como la el campo de la Realidad Virtual. En este sentido, consideramos nuestras aportaciones como un pequeño avance en la dirección correcta para abrir nuevas posibilidades en el estudio de este campo.

## Capítulo 8

# Aportaciones individuales

### 8.1. Sandra Alonso Paz

Los primeros pasos del proyecto fueron la investigación del campo. En mi caso, fue el estudio de la [expresión corporal](#): cómo se medía, qué herramientas existían y cómo se extraían emociones de esta. Durante mi investigación me puse en contacto con distintos proveedores de este tipo de tecnologías como EmteqLabs e Iván García-Magariño, creador de [Emopose](#) y que, gracias a algunos acuerdos, luego formaron parte de CVR-MC. Además, realicé varias pruebas con uno de los posibles dispositivos que podía haber sido utilizado para la lectura del lenguaje no verbal, la [Kinect](#). Por último, me volqué en el estudio de unidades de acción en la expresión corporal y qué combinación de estas hacían expresar una emoción. Esta parte fue clave para la generación del Data Set de posiciones de Classroom VR-Motion Capture.

Como parte de la investigación y en caso de tener que crear nuevas escenas conflictivas, nuestros tutores me pusieron en contacto con la directora del IES Carmen Conde. Me encargué de la organización y colaboré junto con [Andrés Puente Rodríguez](#) y [Antonio Luis Suárez Solís](#) de la transcripción de los [partes](#) que nos proporcionaron. Después de una revisión fueron enviados al equipo de Barcelona.

En cuanto a la gestión del proyecto he podido poner en práctica conceptos de distintas asignaturas que he cursado durante la carrera como “Gestión de proyectos software” e “Ingeniería del Software”. Propuse un modelo basado en [metodologías](#) ágiles en el que fueron clave las reuniones semanales del equipo técnico y la transcripción de las mismas para ver la evolución del proyecto, la propuesta de objetivos semanales y la comunicación horizontal dentro del equipo. Además, también he podido hacer un diseño real de requisitos con el cliente, en este caso el equipo de Barcelona, y mantener el contacto directo con ellos para los posibles cambios de requisitos y sus posteriores validaciones.

En lo referido al experimento realizado en Barcelona, elaboré el [diseño experimental](#) y junto con Andrés Puente, redactamos las preguntas de la [entrevista final](#) que se hicieron a los usuarios al final del experimento. Una vez ya en Barcelona, fui encargada de la primera etapa: [explicación de la prueba](#) y de la supervisión del [desarrollo de la prueba](#) donde tomé nota de las posibles correcciones. Más tarde participé en la redacción del [paper](#) que se incluye en el anexo y que fue enviado al congreso internacional ICCE 2021.



En cuanto a la redacción de la memoria del proyecto, propuse el índice y realicé la estructura general. Además, redacté los apartados del capítulo de [Estado del Arte](#), [Experimentación](#) y en el capítulo de [Metodología y tecnología](#), describí el subapartado de metodología del trabajo. Por otra parte, en el capítulo de [Diseño y desarrollo de la herramienta](#), redacté los subapartados de [Propuesta de ideas e investigación](#), [Incorporación a Didascalia](#), [Investigación sobre el lenguaje corporal y las emociones](#). Además, fui la encargada de hacer los apartados [Motivación](#), y [Resumen](#) junto a su correspondiente traducción. Asimismo, realicé el estudio de los resultados arrojados por las pruebas realizadas en Barcelona, para componer el apartado de [Resultados](#) y [Trabajo Futuro](#). Por último, di formato a la memoria en *LaTeX* e incluí el trabajo de mis compañeros en la misma.

## 8.2. Daniel López Acero

Durante las investigaciones previas al inicio del proyecto mi aportación personal fue un estudio de los distintos tipos de dispositivos que podrían utilizarse para la captura corporal y facial como [Unity](#), [Barracuda](#) o [Xnect](#). Además de ponerme en contacto con los proveedores de estos dispositivos, también pude hacer pruebas con alguno de ellos.

Tras las investigaciones tome un papel más técnico, colaborando en la reestructuración de la aplicación Classroom VR juntos con [Andrés Puente Rodríguez](#). De esta manera, la acomodamos a la nueva estructura que permitiera la generalización de escenas.

A continuación, realicé la transformación del proyecto a RV para permitir al jugador elegir la plataforma de juego en función del hardware disponible. Asimismo, desarrollé la interfaz y su comportamiento y posteriormente, tuve que transformar esta a realidad virtual, para permitir la interacción con ella. Por último, cree un player adaptado a realidad virtual que pudiese interactuar con la interfaz y con el aula sin que se perdieran funcionalidades, esta fue una aportación vital para el proyecto.

En relación con la captura de movimiento, trabajé en la sincronización del traje [Perception Neuron](#) con nuestra herramienta, también colaboré en la adaptación del sistema de captura de emociones tras la incorporación de Emopose al proyecto. Además, experimenté con la herramienta de [EmoPose](#) comprobando su eficacia y validando si era una herramienta factible para incluirla en nuestro proyecto. Tras los buenos resultados, concluí que era perfecta para su inclusión.

En lo que concierne al [experimento](#) realizado en Barcelona, colaboré con la propuesta de preguntas de investigación. Una vez decididas, contribuí en su análisis y en la posterior redacción del [paper](#) incluido en el anexo. Aunque no acudí a realizar la fase de pruebas en Barcelona, colaboré como parte del equipo de soporte desde Madrid.

Por último, respecto a la memoria del proyecto, redacte varios apartados dentro de [Metodología y tecnología](#), como [Tecnologías utilizadas](#), [Dispositivos utilizados](#) y en el capítulo [Diseño y desarrollo de la herramienta](#), [Transformación a VR](#), además de colaborar en múltiples apartados como: [Reestructuración de la aplicación](#), [Incorporación de Emopose](#), [Correcciones](#), [Motivación](#) y [Discusión de resultados](#).

### 8.3. Andrés Puente Rodríguez

El primer paso del proyecto fue la investigación inicial, me centré en aquellas plataformas y [herramientas que detectan el lenguaje corporal](#). A continuación, planteé la estructura de la aplicación. Realicé unos esquemas de la posible estructura y tras discutir con el equipo procedí a implementarla modificando poco a poco la antigua existente para permitir la [generalización](#).

Centrándome en Classroom VR-Motion Capture (CVR-MC), inicialmente realice un estudio de los scripts de generación de las situaciones, para estudiar qué cosas tienen en común. Con esto, procedí a generar los Scriptable Objects que nos permitieran generar las situaciones desde el editor, para facilitar al máximo posible la creación de nuevas situaciones.

El resultado final de los mismos son 3 tipos de scriptable Objects: uno para la generación de la clase y alumnos, otro para las [generalizaciones](#) de las situaciones y un último para las especificaciones de los posibles caminos a tomar en cada situación.

Por otra parte, incorporé la herramienta de [Emopose](#) a una nueva escena, permitiéndonos la captura de poses para asociarlas a emociones que, además, creó el sistema que nos permitió capturar la emoción del usuario cada determinado tiempo durante la simulación y ajustarlo para evitar los fallos por el movimiento lateral del avatar.

Fui responsable de la [división](#) de la aplicación en 2 escenas, una encargada de menús y la otra de la generación de la situación. Estas dos escenas están gestionadas por un GameManager, que también desarrollé. Este, permite la elección de plataforma de juego, por lo que está generalizado para trabajar en ambas. Además, generé las clases en función de la situación, la lógica de simulación y transcurso de las situaciones en un único script que gestiona la escena de juego “MySceneManager”.

Seguidamente, modifiqué la manera en que se utiliza el reconocimiento de voz, para permitir, a cada elección de camino, indicar que palabras requerían atención. Más tarde se juntaron todos los sistemas de reconocimiento y análisis de voz en un único manager. Además realice el CSVSerializer encargado de guardar la información recopilada por la aplicación en los logs.

Mi aportación al proyecto además de técnica, ha sido de coordinación en el desarrollo, decidiendo en qué cuándo comenzar o posponer tareas según la planificación y asignarlas a los diferentes miembros del equipo de desarrollo.

Estuve presente en la realización de la prueba piloto en Barcelona donde se me asignó el rol de “Senior Developer”, realizando cambios y sugerencias que se hacían durante la prueba para mejorar la experiencia de usuario. Además, ayude con las [correcciones](#) finales, tras realizar la prueba y analizar el feedback que se nos proporcionaron los usuarios.

Por último, participé en la transcripción de [partes](#) del Instituto de ESO para su posterior lectura por el equipo de Barcelona.

## 8.4. Antonio Luis Suárez Solís

Durante la fase inicial de investigación mi trabajo se centró en la búsqueda de herramientas de reconocimiento de emociones mediante el análisis de la voz. Más concretamente, cómo capturaban, analizaban y medían las variables relacionadas con la tonalidad. Tras encontrar diversas tecnologías que podrían ser utilizadas para nuestro proyecto, me intenté poner en contacto con *audEERING*, una empresa alemana desarrolladora que provee un plugin de Unity capaz de extraer las emociones de una fuente de audio. Finalmente, no pude mantener el contacto con ellos por lo que no lo incluimos en la versión final del proyecto.

Durante la reestructuración del proyecto, implementé el *GameManager* y sus primeras funciones para gestionar la selección de los packs y el paso entre escenas. Además, implementé el *MenuManager* antes de la inclusión del modo VR, por lo que este podría incluir más funcionalidades en el futuro.

Por otra parte, estuve encargado de revisar el proyecto inicial tras la reestructuración, para volver a incluir los scripts de control del audio que habían quedado fuera del proyecto al reorganizar la escena de juego, como el *MicrophoneManager*. Además, tras la implementación de VR en el proyecto, vimos que el paisaje por defecto de Unity no era lo suficientemente inmersivo, así que investigue los distintos tipos de SkyBox que Unity te permite utilizar, eligiendo finalmente un SkyBox tipo 6-sided Box con texturas obtenidas de Unity Asset Store, llamado [AllSky].

Para finalizar, implementé ciertas correcciones en la aplicación, cuya intención es facilitar el trabajo futuro y aumentar el carácter inmersivo de la aplicación, quedando estas detalladas en el subapartado [Correcciones](#).

En cuanto a otras actividades, participé en la transcripción de partes llevada a cabo en el I.E.S. Carmen Conde junto con [Andrés Puente Rodríguez](#) y [Sandra Alonso Paz](#).

En lo referido a la memoria, redacte el subapartado [Tecnologías y comunicación no verbal](#), el apartado [Objetivos](#), y los subapartados [Diseño Inicial](#), [Transcripción de partes de IES](#), y [Correcciones](#), del apartado [Desarrollo cronológico de la aplicación](#).

## Capítulo 9

# Conclusiones

Para nuestro Trabajo de Fin de Grado, hemos diseñado, desarrollado y evaluado la herramienta Classroom VR-Motion Capture, un entorno de realidad virtual que pretende mejorar los aspectos comunicativos de los profesores de secundaria para la perfección de sus habilidades de gestión de los conflictos dentro del aula. Además, realizamos una prueba piloto para responder a dos principales preguntas de investigación:

- *RQ1: ¿Es factible utilizar el sistema CVR-MC en la formación docente para contribuir al aprendizaje de competencias comunicativas de gestión de clima de aula?*
- *RQ2: ¿Es posible capturar el lenguaje no verbal y las emociones que manifiestan los participantes durante la simulación? ¿Se corresponden con las que manifiestan en un entorno real?*

Durante la [fase de experimentación](#) obtuvimos [resultados](#) muy positivos. El principal, es que, según las apreciaciones de los participantes, la herramienta ha resultado muy útil para entrenar y permitir a los profesores mejorar sus habilidades al resolver situaciones complejas en el aula. Los participantes han destacado la importancia de la inmersividad del sistema y su realismo. Además, todos los participantes han considerado la herramienta como un complemento ideal a introducir en la formación de los profesores noveles.

De esta manera respondemos a la primera pregunta de investigación de manera afirmativa. Sí es un sistema factible para contribuir al aprendizaje de docentes en formación en competencias comunicativas de gestión de clima de aula.

Por otra parte, nuestro sistema permitía inferir las emociones del usuario a través de su expresión corporal mediante un sistema de captura de movimiento. A este respecto, la principal conclusión es que es necesario un estudio más extenso sobre la captura y análisis de la expresión corporal en entornos virtuales. Debido a que los participantes tenían que llevar todos los dispositivos inmersivos necesarios propios de la realidad virtual y la captura de movimiento (gafas, mandos y traje), su expresión corporal se veía comprometida. Observamos que los gestos eran mucho más discretos y escasos que los que se realizan de manera natural en una situación real. De esta manera, nuestro sistema basado en Emopose (una aplicación capaz de inferir emociones a través de la expresión corporal), no era capaz de detectar todas las emociones que los participantes decían sentir.

En este sentido, debemos concluir que la segunda pregunta de investigación no obtuvo un resultado completamente afirmativo. Aunque sí es posible la captura del lenguaje no verbal y estimar la emoción principal del usuario, no podemos hacer un seguimiento de estas posiciones si el equipamiento propio de la realidad virtual no permite a los usuarios mostrar su lenguaje corporal.

En cuanto a la [reestructuración del proyecto](#), consideramos que ha sido clave para poder incluir nuevas situaciones y editar las establecidas de manera sencilla y sin necesidad de indagar en el código de la aplicación. Esto ha facilitado en gran medida el desarrollo del proyecto, y consideramos que podría ser de suma utilidad para las futuras ampliaciones de este proyecto y para los posibles futuros experimentos.

Asimismo, la [integración de Realidad Virtual](#) en la herramienta ha demostrado ser un pilar básico para la inmersión del usuario. Tal y como señalaron los participantes del experimento, este entorno permite al docente gestionar el aula de la manera más fiel posible a como lo haría en una clase real. Si bien la simulación está aún lejos de una realidad total, es un gran paso para conseguir una inversión completa para la práctica de este tipo de situaciones producidas en el aula.

Nos gustaría finalizar estas conclusiones haciendo hincapié en lo realmente importante que es la comunicación no verbal en todo tipo de interacciones sociales y más concretamente en la educación. Como decíamos en la introducción de este documento, contar con docentes formados y experimentados es la base principal para construir nuevas y mejores generaciones. En este sentido, el desarrollo de las habilidades de gestión y mantenimiento de un buen clima de aula son primordiales para asegurar el correcto desarrollo de las lecciones impartidas por los docentes y para construir mejores relaciones profesor-alumno. Para conseguir este propósito es necesario contar con habilidades de comunicación y que estas sean efectivas (que lo que digas con palabras se corresponda a lo que expresas corporalmente). En nuestra opinión este debe ser un requisito fundamental en el programa de formación a nuevos docentes.

## 9.1. Objetivos cumplidos

En cuanto a los [objetivos](#) que planteamos al comienzo del proyecto, hemos logrado completar la gran mayoría de ellos, tanto en lo referido a las implementaciones, como en el experimento y la posterior evaluación de resultados.

El objetivo principal del proyecto era desarrollar un entorno de realidad virtual para que los docentes en formación puedan mejorar sus competencias comunicativas de gestión del clima en el aula en un entorno seguro. Además, este debería contar un sistema de estimación de emociones basadas en el análisis del lenguaje no verbal. Classroom VR-Motion Capture cumple todos estos requisitos.

En cuanto a los [objetivos específicos](#), hemos realizado un buen trabajo alcanzando muchos de los planteados.

En primer lugar, hemos completado la [reestructuración](#) del código de la aplicación que nos permite ampliar el Set de situaciones problemáticas. Con esto hemos sentado la base para futuras ampliaciones del proyecto CVR-MC.

Además, hemos incluido en el sistema la captura de posiciones corporales del docente y hemos desarrollado un sistema de estimación de emociones en base a estas. Aunque no ha ofrecido resultados completamente satisfactorios, consideramos que es un primer paso para alcanzar un análisis completo de las emociones que produce en el docente la gestión de conflictos en el aula.

El tercer objetivo específico cumplido ha sido [integrar el proyecto en realidad virtual completa](#). Además hemos establecido una estrecha relación con el proveedor de [EmteqLabs](#), gracias a ello, se ha acordado obtener un prototipo para la continuación del desarrollo del proyecto en el próximo año.

Por otra parte, se propuso llevar a cabo un experimento para validar la factibilidad y usabilidad de la herramienta. Este objetivo también ha sido cumplido gracias a la [fase de experimentación](#) llevada a cabo en Barcelona con profesionales del sector de la educación. Asimismo, conseguimos [serializar](#) adecuadamente los principales datos de las simulaciones llevadas a cabo en este experimento. Esta base de datos, podrá ser utilizada para el entrenamiento de futuros sistemas.

Por último, con el fin de hacer pública nuestra investigación y el desarrollo de Classroom VR-Motion Capture, escribimos un artículo llamado [Evaluation of a Motion Capture and Virtual Reality Classroom for Secondary School Teacher Training](#). Siguiendo la estructura de esta memoria, incluimos los apartados más importantes del estudio como la introducción de temas a tratar (herramientas tecnológicas para la formación docente y el análisis del lenguaje no verbal mediante tecnologías de captura del movimiento), los objetivos principales, estructura de la aplicación, diseño experimental y sus correspondientes resultados, conclusiones y trabajo futuro para mejorar y ampliar las funcionalidades de CVR-MC. Este artículo fue presentado al congreso *29th International Conference on Computers in Education. Asia - Pacific Society for Computers in Education (ICCE 2021)*

Todos estos objetivos cumplen los requisitos solicitados por el equipo Didascalia (RTI2018-096401-A-I00) y servirán como base para futuras investigaciones y ampliaciones de la herramienta. Por este motivo consideramos que el proyecto ha sido un éxito.

# Capítulo 10

## Conclusions

In this project we show the design, development and evaluation of Classroom VR-Motion Capture application, a virtual reality environment which aims to improve the communicative skills of secondary school teachers in order to improve their conflict management skills. In addition, we conducted an experiment to answer two main research questions:

- *RQ1: Is it possible to use the CVR-MC system in teacher training to improve the communications skills for classroom climate management?*
- *RQ2: Is it possible to capture non-verbal language and their relevant emotions that are expressed by the participants during the simulation? Do they match with those expressed in real environments?*

During the [test](#), we obtained plenty of optimistic [results](#). Firstly, according to the participants' appreciations, the tool will be very useful for inexperienced teachers to improve their skills, concretely, solving complex situations in the classroom. Moreover, participants highlighted the immersivity sense they have experienced while using the system and therefore, its realism. In addition, all participants have considered the tool as an ideal complement to be introduced in new teachers training.

Consequently, we can answer the first research question affirmatively. CVR-MC can contribute in the formative process for teachers in training, specifically in communication skills of classroom climate management

ClassRoom VR-Motion Capture can gauge user's emotions through their corporal expression using a motion capture system. In this regard, the main conclusion we obtained is that more study is needed in capture and analysis of body expression in virtual environments. As the participants had to wear all the necessary immersive devices typical of virtual reality and motion capture (glasses, controllers and suit), their body expression was softer. We observed that their gestures were much more discreet and scarce than those which they performed naturally in a real situation. In this way, our system based on Emopose (an application capable of estimating emotions through corporal expression), was not able to detect every participant's emotions which were felt.

Therefore, we must conclude that the second research question did not obtain a completely affirmative result. Although it is possible to capture non-verbal language and estimate the user's main emotion, we cannot track these positions if the virtual reality equipment does not allow users to show their body language.

Focussing on the [restructuring of the project code](#) has been very useful for including new situations and editing the ones which were already created easily and without investigating the application code. This has greatly facilitated the development of our project, and in our opinion, it could be very useful for future developments and experiments of this project.

On the other hand, Virtual [Reality integration](#) has proven to be indispensable for user immersion. As pointed out by the participants of the experiment, this environment allows the teacher to manage the classroom behaviours as they would do it in a real class. Although the simulation is still far from a reality environment, it is a great step to achieve a complete investment for the practice of this type of situations produced in the classroom.

Finally, We would like to emphasize the importance of non-verbal communication is in every social interaction and especially, in education. As we commented at the introduction, having trained and experienced teachers is the base for building new and better generations. Therefore, the development of management skills and maintenance of a good classroom climate are essential to ensure the correct development of the lessons and to build better teacher-student relationships. To achieve this purpose it is necessary to have effective communication skills (what you say in words are in concordance with what you express with your body). In our opinion, this should be a fundamental requirement in the training program for new teachers.

## 10.1. Accomplishment objectives

Regarding the [initial objectives](#), we have completed the vast majority of these, not just the ones in terms of implementations, but also those related to the experiment and the results evaluation.

The main objective of the project was to develop a virtual reality environment which can contribute in the formative process to improve their communication skills in climate management in the classroom safely. Moreover, it should estimate emotions based on the analysis of non-verbal language. ClassroomVR-Motion Capture meets all these requirements.

Regarding the [specific objectives](#), we have reached plenty of them.

Firstly, we have completed the [restructuring of the application code](#) which allows us to expand the Set of problematic situations. Consequently, we have laid the foundation for future extensions of the CVR-MC project.

Moreover, we have included the teacher's body position capture and we have developed a system for estimating emotions based on these positions. Although it did not obtain completely satisfactory results, we consider it as a first step to achieve a complete analysis of the teachers emotions produced by conflict management in the classroom.



The third specific objective achieved has been the integration in full virtual reality. We have also established a close relationship with [EmteqLabs](#). Consequently, we will obtain a prototype for the future development of the project in the next year.

On the other hand, it was proposed to carry out an [experiment](#) to validate the feasibility and usability of the tool. This objective has also been achieved thanks to the experimentation phase carried out in Barcelona with professionals from the education sector. Likewise, [serialized](#) the main data of the simulations carried out in this experiment. This database can be used for training future systems.

Finally, in order to publish our research and development of Classroom VR-Motion Capture, we wrote an article called *[Evaluation of a Motion Capture and Virtual Reality Classroom for Secondary School Teacher Training](#)*. Following the structure of this report, we include the most important sections of the study such as the Introduction of topics to be discussed (IT tools for teacher training and the analysis of non-verbal language using motion capture technologies), the main objectives, structure of the application, experimental design and the results we have obtained, our conclusions and future work to improve and extend the functionalities of CVR-MC. This article was presented at the *29th International Conference on Computers in Education. Asia - Pacific Society for Computers in Education (ICCE 2021)*

All these objectives meet the requirements requested by the Didascalía team (RTI2018-096401-A-I00) and will be the base for future research and extensions to the tool. For this reason we consider that the project has been a success.

# Capítulo 11

## Trabajo futuro

Los próximos pasos previstos en el desarrollo de la aplicación están enfocados en su mayoría a la experiencia de usuario. Los datos arrojados por el [experimento](#) hacen ver que, aunque los usuarios se sentían cómodos en el entorno, harían pequeñas mejoras en cuanto a su usabilidad.

Uno de los principales problemas que tuvieron los usuarios durante la ejecución entorno virtual fue el manejo de los dispositivos y adecuación al entorno. Un 43 % reflejó en sus sugerencias la posibilidad de hacer una pequeña práctica antes de comenzar con la prueba. Para dar solución a este problema proponemos incluir un escenario inicial similar al de las escenas que se ejecutarán a continuación. De esta manera, una vez que el usuario se sienta cómodo, podrá pasar a las escenas principales habiendo mejorado su experiencia de usuario.

Por otra parte, debemos buscar un método permita que el usuario pueda navegar por la escena sin necesidad de utilizar los mandos propios de la realidad virtual. El principal motivo de este cambio es que, al tener las manos ocupadas, los participantes no pudieron moverse libremente y, por tanto, mostrar de manera natural el lenguaje corporal propio de esa situación.

Además, vemos necesario ampliar el rango de opciones que se dan al usuario a la hora de responder ante la situación crítica ya que algunos señalaron no estar de acuerdo con ninguna de las ofertadas. Esto puede realizarse o bien añadiendo nuevas opciones o bien modificar las ya existentes haciendo que sean más abstractas y por lo tanto menos específicas. De esta manera también permitiremos a los usuarios ser más creativos en sus acciones.

En cuanto a los dispositivos empleados, creemos que sería muy beneficioso mantener el contacto con EmteqLabs y esperar a que terminen el desarrollo de sus [gafas](#). En nuestra opinión, el análisis de otras componentes adicionales como la actividad cerebral puede aportar información muy relevante para este tipo de investigaciones.

Por otra parte, gracias a los resultados estructurados obtenidos por CVR-MC, sería factible utilizar un sistema de redes convolucionales que nos permitiera un análisis más exacto de las posiciones. Podríamos entrenarla con los datos obtenidos y mostrados en el apartado [Resultados obtenidos por CVR-MC](#) e incorporarla al sistema. De esta manera

sería la red la que nos indicará la emoción detectada, en función de los datos que recibe, en lugar de un sistema basado en casos y vecino más cercano como es [EmoPose](#). Esto añadiría fiabilidad a los resultados.

Se propone además implementar nuevos escenarios en otros lugares de la escuela. Los conflictos no solo suceden durante la clase tradicional. Los usuarios proponen que se les presente otro tipo de escenarios como el pasillo de la escuela u otras clases con distinta disposición, como puede ser el aula de tecnología o de artes.

Por último y más concretamente en cuanto a próximos experimentos, proponemos que, durante los experimentos, además de grabar toda la ejecución con una cámara externa, también guardemos el gameplay. Esto permitirá a los investigadores hacer un mejor análisis de las acciones tomadas por el usuario.

## Capítulo 12

### Anexo

#### 12.1. Publicación de un paper en el ICCE 2021

# Evaluation of a Motion Capture and Virtual Reality Classroom for Secondary School Teacher Training.

Sandra ALONSO <sup>a</sup>, Daniel LÓPEZ <sup>a</sup>, Andrés PUENTE <sup>a</sup>, Alejandro ROMERO <sup>a</sup>,  
Ibis ALVAREZ <sup>b</sup> & Borja MANERO <sup>a</sup>

<sup>a</sup> *Universidad Complutense de Madrid, Spain*

<sup>b</sup> *Universitat Autònoma de Barcelona, Spain*

[sanalo05@ucm.es](mailto:sanalo05@ucm.es)

[danilo08@ucm.es](mailto:danilo08@ucm.es)

[andpuent@ucm.es](mailto:andpuent@ucm.es)

[alerom02@ucm.es](mailto:alerom02@ucm.es)

[ibismarlene@gmail.com](mailto:ibismarlene@gmail.com)

[bmanero@ucm.es](mailto:bmanero@ucm.es)

**Abstract:** Nowadays having qualified and experienced teachers in school classrooms is considered to be of the highest priority in any society. However, most teachers report that they haven't received sufficient practical training to manage disruptive situations in the classroom. Fortunately, virtual reality can provide a solution to this issue.

This paper will introduce ClassroomVR-MotionCapture (CVR-MC) and its evaluation by some experts. CVR-MC is an IT tool that can simulate a virtual classroom. This allows different users to face real-life problems that usually take place in real classrooms. The system captures the tone of voice, the substance of the speech, the gaze and corporal movements of users. Virtual students will react according to these parameters.

To evaluate the usability and functionality of the tool, we organized a study in which 14 education professionals took part. The research questions that guided the study were the following:

1. Is it possible to use the CVR-MC system in teacher training in order to improve teachers' communication skills for classroom climate management?
2. Is it possible to capture non-verbal language and related emotions that are expressed by the participants during the simulation? Do they match those expressed in real environments?

Finally, the main conclusion of our study is that many participants described the CVR-MC system as a friendly, safe and feasible environment for teacher training, especially for improving their classroom climate management competence. However, the study also found that the detection of emotions through the users' body expression did not match the emotions they reported feeling during the test.

**Keywords:** education, non-verbal language, teaching, emotion recognition, virtual reality, serious games.

## 1. Introduction

The ICCE From an ecological perspective of teaching, classroom management creates, through the actions of teachers, an environment that facilitates achieving the learning goals and improving the students' socio-emotional well-being (Emmer & Stough, 2001). However, despite teachers' efforts to create and maintain a productive learning environment, unexpected events may occur in the classroom that could interfere with instructional activities. The success of classroom management is dictated by the teacher's competence in understanding and interpreting conflictive events, in immediate circumstances.

Conflicts threaten interpersonal relationships and may lead to destructive events that can dramatically affect teaching and learning conditions. Most of the emotions that teachers experience are influenced by the situations that arise in the classroom (Sutton et al., 2009).

Due to the importance of this issue, there is an urgent need to overcome teacher training deficiencies by improving teachers' competence in classroom climate management. The Organization for Economic Co-operation and Development (OECD) and Teaching and Learning International Survey (TALIS) (Schleicher, 2020) state: "Spain is well below the average for teachers who refer having received training regarding management of students, conflicting behavior and classroom climate. Less than half of Spanish teachers (40%) do feel prepared to control a class. This study also reveals that Spanish teachers exceed in their responses at maintaining order in class".

Based on the deficiencies recognized in current teacher training programs, more specifically, in relation to the competence for classroom climate management, our objective is to explore the usefulness of a virtual classroom system that has been developed with the intention of promoting the learning of this competence in initial teacher training.

### *1.1 IT tools for teacher training*

Our work is not the first approach to developing tools for improving teaching skills using virtual reality environments and technologies. Firstly, a US team developed the TeachLiveETM system (Barmaki & Hughes, 2015), a tool for analyzing non-verbal language. The authors conducted an experiment where only half of the participants received feedback on their non-verbal language and how to improve it during their lessons. They concluded that users who were given feedback experienced a significant improvement compared to those who did not receive it. Although this tool seems to have had good results, it did not consider the different elements of non-verbal language, such as gaze direction or voice tonality.

Another example was implemented in Japan (Huang et al., 2016). The researchers developed a system for training new non-experienced teachers. This environment simulated a high school class and used a Kinect device that registered the movement of the head and hands, as well as the user's voice.

A more recent approach was the study carried out by researchers from the Swedish FHNW School of Education (Wyss et al., 2021). They developed a system that allows improving the professional vision of future teachers thanks to eye tracking and post-hoc think-aloud verbalizations.

It is also worth mentioning the German Knowledge Media Research Center (Leibniz-Institut für Wissensmedien) study (Sümer et al., 2021), where the authors validated a system that permits the evaluation of students' attention span through visible indicators of their level of participation in learning. Although the system has proven to be efficient, it could be more effective using automated analysis.

All of these systems provide very useful information about how non-verbal communication affects the interactions with students. The information collected by these studies support the view that the systems are appropriate for improving competencies of non-experienced teachers. Moreover, the system also allows teachers in training to have as many goes as needed to feel in command of the situation, which is the ultimate goal of the training exercise. However, we have not found any tool that can combine the detection and evaluation of non-verbal behavior (emotions and attitudes) with strategies to manage classroom conflicts.

Considering the close relationship between both fields, this research tries to provide a VR environment (which we have called CVR-MC) that allows future teachers to develop their competence in managing classroom climate. We test users' competence by exposing them to three conflictive situations that are quite frequent in high school classrooms.

The tool in which our work is based is called ClassroomVR (Bocos Corredor et al., 2020). This system generates different conflictive situations that must be resolved by the teachers using the resolution methodologies suggested by the system.

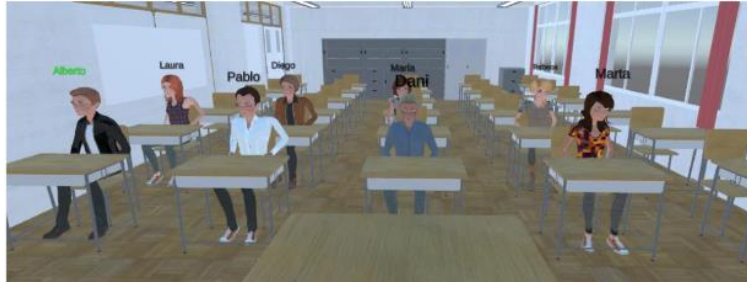


Figure 1. ClassroomVR environment.

### 1.2 Non-verbal analysis using motion capture

Several years have been needed for studying non-verbal language and how to capture it due to the complexity of the field. During this time, different theories and methods of analysis have been proposed (Viera, 2015). In a real situation, it is possible to analyze people's non-verbal language based on their natural movements (Torres, 2019).

By observing body language, we can estimate the emotions people may be feeling (Ruano Arriagada, 2004). These studies use a Kinect device for detecting emotions through the user's body expression. They all agree that it is difficult to determine the predominant emotion in the users just based on the person's posture. It has been claimed that a complete psychological study is needed.

Another proposed tool is Emopose (García-Magariño et al., 2019), which is able to analyze an image projected on a 3D avatar and recognize one of the 6 basic emotions through the user's posture (anger, disgust, fear, happiness, sadness and surprise). This tool is based on the closest neighbor algorithm, choosing from its database the position closest to the one analyzed.

### 1.3 Objectives and structure

The main objective of this paper is to show our tool validation. This study seeks to answer these research questions:

- RQ1: Is it possible to use the CVR-MC system in teacher training to improve the communications skills for classroom climate management?
- RQ2: Is it possible to capture non-verbal language and their relevant emotions that are expressed by the participants during the simulation? Do they match with those expressed in real environments?

This paper is structured in six different sections: the first one reviews the existing literature. The second one explains the application's architecture and its functionalities. Section three details the experimental design. The fourth section presents the result of the evaluation of the CVR-CM tool. Section five presents the discussion, and finally, section six shows the conclusions and future work.

## 2. CVR-MC Architecture

The work by (Bocos Corredor et al., 2020) established the basis of CVR-MC aims to practice and improve teaching skills. Consequently, the authors created an environment that allows teachers to experience the climate and emotional factors present in the classroom. Based on this application, several extensions have been developed and they are all presented in Figure 2.

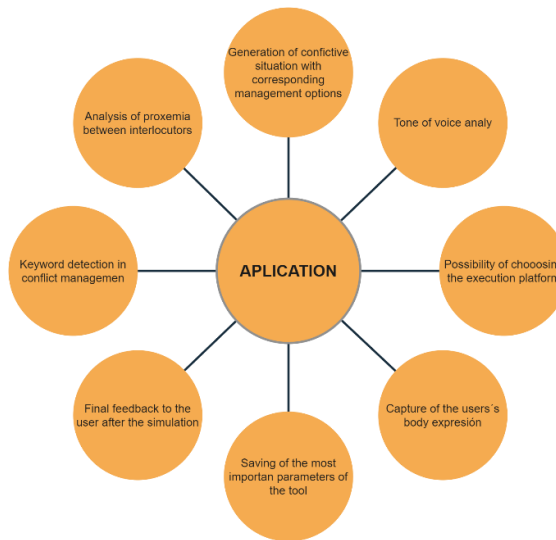


Figure 2. Architecture of CVR-MC

The developed functionalities are detailed below:

- **Generation of conflictive situations and their corresponding management options:** the tool generates a conflictive situation based on generalized previous parameters with their corresponding management options. This allows situations to escalate and allows them to be easily generated.
- **Tone of voice analysis:** we analyze the teacher's tone of voice to detect changes during key moments of the simulation.
- **Analysis of proxemia between interlocutors:** the tool analyzes the distance between two interlocutors (student and teacher) in order to determine what the simulation response will be.
- **Keyword detection in conflict management:** the words used by the user are recorded. These are classified to determine the response to the simulation and relate them to the user's emotions.
- **Possibility of choosing the execution platform:** CVR-MC allows you to choose the execution platform; this can be a virtual reality environment or a desktop environment (PC).
- **Analysis of the user's body expression to establish the predominant emotion:** taking advantage of the use of the Perception Neuron and EmoPose system (García-Magariño et al., 2019), we captured and analyzed the teacher's body expression in order to associate it with an emotion that could be transmitted to the students without being aware of it.
- **Storage of the most relevant parameters of the simulation:** during simulation those parameters that are considered most important are saved for analysis. This is done with the tone of voice, language, keywords, chosen path and emotion associated with the words and body language.
- **Final feedback to the user after the simulation:** CVR-MC analyzes the data captured during the course of the simulation to give a final feedback to the user about his/her behavior and intention perceived during the simulation.

### 3. Evaluation of CVR-MC.

#### 3.1 Participants

Fourteen participants, all experienced in the education field, were recruited to take part in the first CVR-M test. Among them were high school principals, teachers and school counselors; teachers of the Master's Degree in Training for Primary or Secondary Teachers and students of the Master's Degree in Teaching.

In terms of teaching experience, the vast majority (57.1%) had been working in education for more than five years. Only two participants (14.3%) had less than two years' experience. With regards to previous experience with virtual reality, 64.3% were already familiar with the environment. The rest admitted to never having tried a virtual reality environment.



### *3.2 Experimental design*

The test took place in the Multipurpose Room of the Faculty of Education of the Universidad Autónoma de Barcelona. The main characteristics of this classroom is its open space and movable chairs. The 14 participants were summoned throughout the day with 30 minutes difference. The test took 25 minutes to complete and consisted of four stages:

#### *3.2.1 Test explanation*

Users were welcomed by a member of the technical team and provided with and asked to sign the consent document, which gave us consent to record audio and video for educational purposes. The candidate was given an ID, which was then used to identify them throughout the whole process. After taking note of some control questions about their previous experiences (teaching or experience with similar tools), they were given tips for the test.

#### *3.2.2 Test*

At this step, where all the material for virtual reality immersion was placed, the user was equipped with Perception Neuron and, afterward, we calibrated it to obtain the most accurate results possible. Then, the Oculus Rift S were adjusted to their head and finally, a headset was placed into one of their ears to allow them to hear what was said during the game and the instructions of the technical staff, with the entire process being recorded.

Before the experience started, we let users walk around the environment and get familiar with it. By pressing the “Play” button, the three situations that the user would face were shown. When the user started the practice run of one of these scenes, a brief description of the environment was shown and the user was expected to simulate an explanation as if they were in a real classroom. After a short period of time, a disruptive situation would occur and the possible options that the user could choose to perform were shown. Afterwards, the player had to simulate one of these actions in a natural way. The game recognized one of the paths taken and showed a final feedback.

Feedback consisted of three screens:

1. It indicated whether the action chosen was the most appropriate one.
2. It showed the main emotion captured by the system during the scene. (fear, anger, joy, disgust, surprise or sadness).
3. It reflected the variation in the tone of the player's voice during the interaction with the virtual students.

Once the three scenes were completed, the devices were removed and the test recording stopped.

#### *3.2.3 Questionnaire post-test*

On the third step, we offered a questionnaire based on the Technology Acceptance Model (TAM, (Davis, 1989)). These questionnaires have been widely used in many educational researches. We adapted the questionnaire proposed by (Huang et al., 2016), who conducted a similar study in Japan. Eighteen items grouped into three dimensions defined by the TAM were included.

- Perceived usefulness: how the user (teacher/teacher in training) perceives that the use of IGLE can improve their competence to manage classroom conflicts (8 items). For example: "I perceived that my behavior - movements, attitudes, words... impacted students in the virtual environment."
- Perceived ease of use: if the user (teacher/teacher in training) perceives that the use of the IGLE will not involve any effort for them'
- Attitude towards app use: refers to the emotions (positive or negative) experienced by the user (teacher/teacher in training) in their experience in the IGLE.
- Behavioral intention: how likely is that the user will use IGLE as an environment to teach and learn about classroom climate management.

To illustrate user answers, we used a seven-level Likert scale with the following correspondence: from total disagreement (1), to total agreement (7). The questionnaire was administered online at the end of the test. This questionnaire is included at the end of this paper.

### 3.2.4 Final interview

The last stage consisted of a semi-structured interview (see next section) guided by a member of the research team. The interview lasted approximately 15 minutes and it explored the evaluation of the experience by the participants, specifically on the feedback provided about their emotions during the execution of the system test. Furthermore, suggestions were written down for possible improvements to the tool.

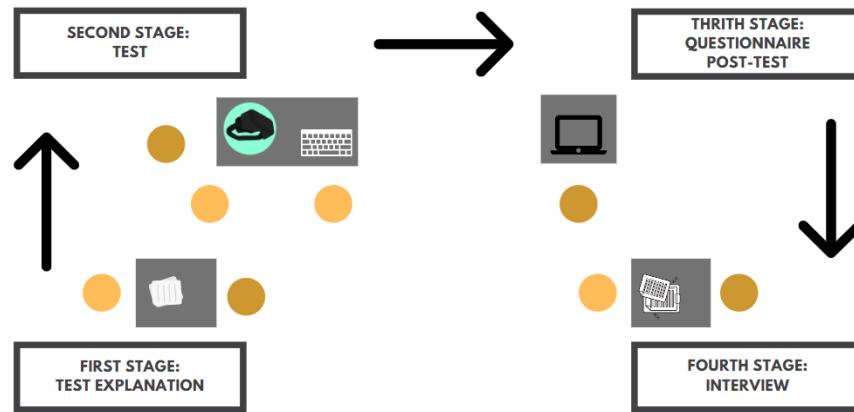


Figure 3. Shows the entire procedure for the test.

## 4. Result

### 4.1 Quantitative result of the test

According to the model used to create the test (Davis, 1989), there are two main variables that affect users' acceptance and adoption of new technologies: perceived ease of use (1) and perceived utility (2). The answers for the questions that made up the post-test questionnaire (based on TAM), which users filed after CVR-MC test, are included below.

Table 1: Summary of the results obtained in the post-test-questionnaire to assess the adaptation of user to CVR-MC (seven-level Likert-type scale)

Criteria	Average	SD
Perceived usefulness	4.7	1.3
Perceived ease of use	5.1	1.3
Attitude towards the use of the IGLE	6.1	0.7
Behavioral intention	6.3	0.8

- **Perceived usefulness**

- The students in the environment have human appearance and attitudes.
- The students' attitudes resemble those found in a natural conflict.
- The number of students is adequate for the exercises that are proposed.

- d. The students in the virtual environment responded appropriately to my actions.
- e. I perceived that my behavior (movements, attitudes, words...) impacted the students in the virtual environment.
- f. I think the students in the virtual environment understood what I said.
- g. Emotions were aroused in me towards the students in the virtual environment. For example: anger, fear, joy, tension...
- h. I was able to adapt my way of managing the conflict according to specific attitudes expressed by the students in the virtual environment.
- **Perceived ease of use**
  - i. I felt like I was in front of a real class.
  - j. The tool was easy for me to use, the environment is easy to understand.
  - k. I felt comfortable using the virtual reality devices.
  - l. The devices (glasses and suit) did not interfere with the actions I wanted to perform.
- **Attitude towards the use of the IGLE**
  - m. I have enjoyed using this virtual learning environment.
  - n. It's fun to practice the management of the classroom climate in this virtual environment.
  - o. I believe that the virtual simulation of disruptive behaviors in the classroom
- **Behavioral intention**
  - p. I believe that the virtual simulation of disruptive behaviors in the classroom helps the learner to improve their ability to handle incidents that arise during classes.
  - q. My teaching competence in managing the classroom climate can improve with the continuous use of this virtual learning environment.
  - r. It would be great to have this virtual learning environment as a professional training tool at school (continuing teacher training) or at university (initial training).

Table 2: Summary of Cronbach's Alpha obtained after the post-test-questionnaire.

Question group	Number of elements	Cronbach's Alpha	Alpha based on standard elements
PU	8	0,815	0,813
UF	4	0,840	0,864
UA	2	0,802	0,841
IU	3	0,617	0,689

Questions from **a** to **h** showed us information about the usefulness of CVR-MC, according to users. For this analysis, we added the responses obtained. In order to answer this question, a Cronbach's alpha was carried out, obtaining as a result an alpha of 0.915 and an average of 4.7 in responses. These responses tell us that users considered that the tool could improve their competence in managing conflicts in the classroom. This matches the great reception shown by the users in the final interview and the argument asking to obtain a safe environment for the practice of these skills.

On the other hand, questions **i**, **j**, **k** and **l** corresponded to perceived ease of use. To analyze the answers to these questions, a Cronbach's alpha was recreated with the sum of all of them. As a result, we obtained an average of 5.1, showing that a large part of the users didn't have to make a big effort to use the tool; however, some felt some discomfort. In the interviews, we were able to find out the limitations detected by the participants.

Next, we analyzed the questions from **m** to **o**, which showed information about the attitude towards use. To analyze the results, we used a Cronbach's alpha that resulted in 0.802 and an average result of 6.1. These results showed that the experience with the tool was relatively good. With this data we observe that although virtual reality environments are something new in this sector, they are also exciting and striking.

Finally, questions **p**, **q** and **r** referred to the intention of using the tool by users. By means of a Cronbach's alpha we analyzed the results of these questions: we obtained an alpha of 0.617 and an average of 6.3. With this and backed up by the interviews, we can conclude that, with certain improvements, the tool could play a part as a complementary use in teacher training.

## 5. Discussion

*RQ1: Is it possible to use the CVR-MC system in teacher training in order to improve teachers' communication skills for classroom climate management?*

Yes, after analyzing the results we can affirm that our tool has potential to be used as a learning system that can contribute to teacher training. Users agreed that CVR-MC can be a highly recommended support tool (4.7 on average) for teachers in training. CVR-MC will allow them to safely practice the necessary techniques to face a conflictive situation. In consequence, it will allow a greater development in their conflict resolution skills.

According to experts' comments, the feedback provided by CVR-MC is highly beneficial for trainee teachers (6.3 on average). Similarly, a joint discussion during teacher training can be very constructive and interesting for future teachers. However, we observed how some participants referred to learning problems with the tool due to manageability, from which we can deduce that training prior to practical use would be helpful.

*RQ2: Is it possible to capture non-verbal language and related emotions that are expressed by the participants during the simulation? Do they match with those expressed in real environments?*

No. Users did not move as initially expected and many emotions were not detected. As a result, the association of emotions to the positions adopted by users did not achieve the expected result. Although the motion capture and emotion recognition system works perfectly in a theoretical environment, in practice we could observe that users felt inhibited when reacting to situations in a virtual environment. This could be due to either the limited time available to them to make a decision or to the type of options to manage the conflictive situation that are presented to them.

At this point, it should be clarified that the equipment worn by the participants (glasses, controls, and motion capture suit) can make the gestures that are produced less significant than those that are made without all these devices. This has meant that the emotions detected by the system (based on Emopose (García-Magariño et al., 2019), a system not based on virtual environments) have not been those that the participants claimed to have experienced. It is therefore necessary to study body movement in virtual environments, since it seems to be different from that used in real environments.

## 6. Conclusions and future work

This document presents the ClassroomVR-MotionCapture virtual reality tool that aims to improve the communication skills of secondary school teachers in order to improve conflict management within the classroom. In addition, the tool has been evaluated by experts in secondary education.

The main result we have obtained is that our tool, as confirmed by users, could be very useful in teacher training, as it would allow teachers to improve their skills in solving complex situations in the classroom. Participants highlighted the importance of the system's immersivity and its realism. Moreover, all participants recognized that the tool is a perfect complement to be introduced in the training of new teachers.

Our system sought to extract users' emotions through their corporal expression with a motion capture system. However, the main conclusion is that it is necessary to study how to track body expression in virtual environments. We concluded that, as the participants had to wear all the necessary devices to enter virtual reality (glasses, controllers and suit), their body expression changed. Consequently, their gestures were much more discreet than those done in real environments. In this sense, our system based on Emopose (García-Magariño et al., 2019) (an application capable of inferring emotions through corporal expression) was not prepared for detecting the emotions felt by participants.

As future work, the first step will be to carry on with the study of body movement within virtual reality. Therefore, we will be able to detect real emotions in virtual environments.

In addition, we will improve user experience by following some of the participants' suggestions. For instance, in to improve the immersion we will replace the initial and final explanatory posters of

the situation with a recorded audio. Finally, we must look for a method that would allow the user to move around the scene without using the virtual reality devices. This is mainly because having their hands occupied showed that participants were not able to move freely and therefore, naturally display the body language expected in that situation.

## 7. Acknowledgements

This project has been funded by the Ministry of Science, Innovation and Universities of Spain (Didascalías, RTI2018-096401-A-I00).

## 8. Annex

Table 3: Average and Standard Desviation of the answers from the test

Question	Average	SD
Do the students in the virtual environment look and behave as real persons?	4.57	1.39
Is the reaction of the students closer to the student with a disruptive attitude natural?	4.07	1.43
The number of students is adequate for the exercises that are proposed.	5.6	1.17
The students in the virtual environment responded appropriately to my actions.	4.8	1.3
I perceived that my behavior (movements, attitudes, words...) impacted the students in the virtual environment.	4	1.41
I think the students in the virtual environment understood what I said.	4	1.25
Emotions were aroused in me towards the students in the virtual environment. For example: anger, fear, joy, tension...	4.92	1.38
I was able to adapt my way of managing the conflict in correspondence to specific attitudes expressed by the students in the virtual environment.	5.28	1.09
I felt like I was in front of a real class.	5.21	1.26
The tool was easy for me to use, the environment is understandable.	5.14	1.05
I felt comfortable using the virtual reality devices.	4.85	1.3
The devices (glasses and suit) have not interfered with the actions I wanted to perform.	5.21	1.73
I have enjoyed using this virtual learning environment.	5.92	0.88

It's fun to practice the management of the classroom climate in this virtual environment.	6.28	0.58
I believe that the virtual simulation of disruptive behaviors in the classroom helps the learner to improve their ability to handle incidents that arise during classes.	6.28	0.69
My teaching competence in managing the classroom climate can improve with the continuous use of this virtual learning environment.	5.93	1.33
It would be great to have this virtual learning environment as a professional training tool at school (continuing teacher training) or at university (initial training).	6.78	0.41

## 9. References

- Barmaki, R., & Hughes, C. E. (2015). Providing real-time feedback for student teachers in a virtual rehearsal environment. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 531–537.
- Bocos Corredor, M., López García, Á., & Díaz Nieto, A. (2020). Classroom VR: a VR game to improve communication skills in secondary-school teachers.
- Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 319–340.
- Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist*, 36(2), 103–112.
- García-Magariño, I., Cerezo, E., Plaza, I., & Chittaro, L. (2019). A mobile application to report and detect 3D body emotional poses. *Expert Systems with Applications*, 122, 207–216.
- Huang, H.-H., Ida, Y., Yamaguchi, K., & Kawagoe, K. (2016). Development of a virtual classroom for high school teacher training. *International Conference on Intelligent Virtual Agents*, 489–493.
- Ruano Arriagada, M. T. (2004). La influencia de la expresión corporal sobre las emociones: un estudio experimental. *Ciencias*.
- Schleicher, A. (2020). Teaching and Learning International Survey (TALIS) 2018. Insights and Interpretations. OECD.
- Sümer, Ö., Goldberg, P., D'Mello, S., Gerjets, P., Trautwein, U., & Kasneci, E. (2021). Multimodal Engagement Analysis from Facial Videos in the Classroom. *ArXiv Preprint ArXiv:2101.04215*.
- Sutton, R. E., Mudrey-Camino, R., & Knight, C. C. (2009). Teachers' emotion regulation and classroom management. *Theory into Practice*, 48(2), 130–137.
- Torres, A. J. (2019). Interacción didáctica y lenguaje no verbal. *Interconectando Saberes*, 7.
- Viera, L. L. (2015). La influencia del lenguaje no verbal en la mediación. *Revista de Mediación*, 8(2), 2340–9754.
- Wyss, C., Rosenberger, K., & Bühner, W. (2021). Student teachers' and teacher educators' professional vision: Findings from an eye tracking study. *Educational Psychology Review*, 33, 91–107.

## 12.2. Consentimiento firmado

**TÍTULO DEL ESTUDIO:** Comunicación efectiva para profesorado de secundaria mediante realidad virtual

**ACRÓNIMO:** Didascalía

Referencia administrativa: RTI2018-096401-A-I00.

Progr

ama Estatal de I+D+i Orientada a los Retos de la Sociedad. MEC. España

**INVESTIGADORES RESPONSABLES**

Borja Manero Iglesias

Instituto de Tecnología del Conocimiento - ITC

Universidad Complutense de Madrid

Correo electrónico: bmanero@ucm.es

Ibis M. Alvarez Valdivia

Dpto. Psicología Evolutiva y de la Educación

Universitat Autònoma de Barcelona

Correo electrónico: ibismarlene.alvarez@uab.es

**PROPÓSITO DEL ESTUDIO**

Considerando vuestra experiencia educativa os invitamos a colaborar en nuestra investigación. Antes de decidir participar en este estudio, es importante que comprenda por qué se está realizando la investigación y qué implicará. Por favor lea la siguiente información cuidadosamente. Pregúntele al investigador si hay algo que no está claro o si necesita más información.

*El objetivo del proyecto Didascalía es enseñar competencias de comunicación a los profesores de secundaria a través de un entorno de realidad virtual, para ayudarles a gestionar el aula y a resolver conflictos. Para ello, se creará un equipo transdisciplinar con experiencia en los siguientes campos: videojuegos educativos, pedagogía, computación afectiva y realidad virtual. El resultado de esta investigación será un nuevo método de aprendizaje basado en un **Entorno de Aprendizaje Gamificado Inmersivo (IGLE, por sus siglas en inglés)**. En este entorno de realidad virtual se enseñará al profesor a manejar una clase utilizando sus competencias comunicativas. El comportamiento de la clase se modelará en base a las observaciones reales que los alumnos de los másteres de formación del profesorado han recogido en sus memorias de prácticas o a través de partes de disciplina. En el entorno virtual, se recogerá información sobre el jugador/docente - medidas biométricas, análisis de voz, movimientos corporales y contenido del discurso, entre otros - para valorar su capacidad de comunicación. La clase virtual reaccionará en base al análisis de los parámetros recogidos, lo que le proporcionará al jugador una realimentación en tiempo real de su actuación. Esto servirá de antesala a las prácticas frente a las clases reales discurso, entre otros - para valorar su capacidad de comunicación.*

**PROCEDIMIENTO**

La metodología propuesta para el proyecto Didascalía se basa en cuatro puntos fundamentales:

1. **Fase de investigación:** recopilación y análisis de situaciones conflictivas más comunes en las aulas de secundaria y las estrategias que utilizan los docentes para gestionarlas. El análisis de estas situaciones es decisivo para identificar actuaciones de estudiantes y de docentes durante estas situaciones comunicativas que posteriormente serán modeladas en el Entorno de Aprendizaje Gamificado Inmersivo (IGLE, por sus siglas en inglés).



2. **Fase de especificación y diseño:** diseño de un entorno de aprendizaje inmersivo en realidad virtual, aplicando técnicas de gamificación. De nuevo, estará centrado, por un lado, en el docente y por otro, en la audiencia.
3. **Fase de desarrollo:** desarrollo del IGLE y evaluación de expertos. Tutores de prácticas de los másteres de formación del profesorado de la Universidad Complutense de Madrid, de la Universidad Autónoma de Madrid y de la Universitat Autònoma de Barcelona evaluarán la herramienta.
4. **Fase de experimentación:** experimentación con usuarios potenciales: profesores de secundaria en formación y en ejercicio. Se recogerán datos de estos experimentos y se analizarán posteriormente. Se busca demostrar que las metodologías desarrolladas son eficaces en la solución de los problemas detectados.

## **COLABORACIÓN QUE SE SOLICITA**

**En vuestro caso, la colaboración que se solicita se circunscribe a la segunda fase. Concretamente os solicitamos participar en la Prueba de usabilidad del prototipo de IGLE que se ha diseñado en base a los resultados de la primera fase de la investigación.**

## **BENEFICIOS**

Participar en el estudio no tiene ningún costo. Ni usted, ni otra persona involucrada en el estudio, recibirá beneficios políticos, económicos o laborales como compensación por su participación.

## **CONFIDENCIALIDAD**

La información que se recoja será confidencial y no se usará para ningún otro propósito fuera de los de esta investigación. **Los investigadores asegurarán el anonimato mediante la asignación de nombres ficticios o códigos** a centros, profesorado o alumnado que pudiera aparecer en la información aportada.

## **VOLUNTARIEDAD DE PARTICIPACIÓN**

Su participación será completamente voluntaria y tendrá el derecho de retirarse en cualquier momento del estudio si usted así lo desea. Igualmente, si en algún momento desea que la información que usted brinda no sea utilizada por los investigadores, lo podrá comunicar y respetaremos su decisión.

Le informaremos de los resultados obtenidos en el estudio. También podrá contactar en cualquier momento al personal de la investigación e informarnos cualquier situación anormal o inesperada.

Después de firmar el formulario de consentimiento, aún puede retirarse en cualquier momento y sin dar una razón. Retirarse de este estudio no afectará la relación que tenga, si la hubiera, con el investigador. Si se retira del estudio antes de que se complete la recopilación de datos, sus datos le serán devueltos o destruidos.

## **CONSENTIMIENTO**

- 
- He leído la información sobre el proyecto de investigación y he tenido la oportunidad de hacer preguntas, las cuales se me han respondido satisfactoriamente.
  - Estoy de acuerdo en participar y he recibido una copia de este consentimiento.

Firma del/la participante \_\_\_\_\_ Fecha \_\_\_\_\_

Firma del/la investigador/a \_\_\_\_\_ Fecha \_\_\_\_\_

### **12.3. Preguntas del cuestionario post-test**

El cuestionario proporcionado en la tercera etapa de la fase de experimentación se puede consultar en el siguiente [Cuestionario post-test].

# Bibliografía

- [Vallejo, 2005] Vallejo, A. P. (2005). La gestión de conflictos en el aula. Factores determinantes y propuestas de intervención. Pérez, M., & Rivera, A.(2005). La orientación escolar en centros educativos, 319.
- [Prats et al., 2005] Prats, J., Raventós, F., Gasòliba, E., Cowen, R., Creemers, B. P., Gauthier, P. L., ... & Standaert, R. (2005). Los sistemas educativos europeos; Crisis o transformación?. Colección de estudios sociales, (18).
- [Rebolledo Gámez, 2015] Rebolledo Gámez, T. (2015). La formación inicial del profesorado de educación primaria y secundaria en Alemania, España, Finlandia, Francia y Reino Unido. Estudio Comparado. Revista Española de Educación Comparada, 25, 129-148.
- [Beraza & Cerdeiriña, 2011] Beraza, M. Á. Z., & Cerdeiriña, M. A. Z. (2011). La formación del profesorado de Educación Infantil. Los docentes, conciencia educativa de la sociedad, 103.
- [Pérez-de-Guzmán et al., 2011] Pérez-de-Guzmán, V., Muñoz, L. V. A., & Vergara, M. V. (2011). Resolución de conflictos en las aulas: un análisis desde la Investigación-Acción. Pedagogía social: revista interuniversitaria, (18), 99-114.
- [Muñoz, 2009] Muñoz, J. M. E. (2009). La formación del profesorado de Educación Secundaria: contenidos y aprendizajes docentes Secondary Education Teacher Training: teaching content and teacher's learning. Revista de educación, 350, 79.
- [ABC] Más de la mitad de los profesores españoles no llega formado en el contenido, la pedagogía y la práctica de la materia que imparte: <https://www.abc.es/>
- [Real Academia Española] Definición de comunicación: <https://dle.rae.es/>
- [Kendon, 1972] Kendon, A. (1972). Review of Birdwhistell's Kinesics and Context. American Journal of Psychology, 83, 441-455.
- [García, 2000] García, I. (2000). Cinética y proxemia en el aula. Educación, 9(18), 157-181.
- [Schmidt, 2013] Schmidt, S. (2013). Proxémica y comunicación intercultural: la comunicación no verbal en la enseñanza de E/LE. Universitat Autònoma de Barcelona,.
- [Poyatos, 1970] Poyatos, F. (1970). Paralingüística y kinésica: para una teoría del sistema comunicativo en el hablante español. In Actas del Tercer Congreso Internacional de Hispanistas (pp. 725-738). El Colegio de México.

- [Botella et al., 2006] Botella, C., García-Palacios, A., Quero, S., Baños, R. M., Bretón-López, J. M. (2006). Realidad virtual y tratamientos psicológicos: Una revisión. *Psicología Conductual*, 3, 491–510.
- [Delgado Pardo & Moreno García, 2012] Delgado Pardo, G., Moreno García, I. (2012). Aplicaciones de la Realidad Virtual en el Trastorno por Déficit de Atención con Hiperactividad: Una aproximación. *Anuario de Psicología Clínica y de La Salud*, 8 (2012), 31-39.
- [Ebert et al. 2014] Ebert, L. C., Nguyen, T. T., Breitbeck, R., Braun, M., Thali, M. J., Ross, S. (2014). The forensic holodeck: an immersive display for forensic crime scene reconstructions. *Forensic Science, Medicine, and Pathology*, 10(4), 623–626.
- [Escartín, 2000] Escartín, E. R. (2000). La realidad virtual, una tecnología educativa a nuestro alcance. *Pixel-Bit. Revista de Medios y Educación*, 15, 5-21.
- [Mendoza, 2016] Mendoza, L. I. U. (2016). Uso de la realidad virtual, en la educación del futuro en centros educativos del Ecuador. *Journal of Science and Research: Revista Ciencia e Investigación*, 1(4), 26–30.
- [RAMALLAL & MURILLO, n.d.] RAMALLAL, D. R. P. M., MURILLO, D. R. A. M. (n.d.). REALIDAD VIRTUAL EN LA RECREACIÓN DE ESCENAS DE CRIMEN COMO COMPLEMENTO FORMATIVO EN LOS ESTUDIOS DE DERECHO CRIMINALÍSTICO (EN TIEMPOS DE COVID-19). UN ESTADO DEL ARTE. *Preferencias Musicales y Consumos de Medios Entre Jóvenes Del Curso de Pedagogía*.
- [Rizzo et al., 2002] Rizzo, A. A., Bowerly, T., Buckwalter, J. G., Schultheis, M., Matheis, R., Shahabi, C., Neumann, U., Kim, L., Sharifzadeh, M. (2002). Virtual environments for the assessment of attention and memory processes: the virtual classroom and office. *Proceedings of the Fourth ICDVRAT*, 3–12.
- [Roehl, 1996] Roehl, B. (1996). Special edition using VRML. Mc Millan Computer Publishers.
- [Valencia Castro et al., 2019] Valencia Castro, J. L., Tapia Vallejo, S., Olivares Olivares, S. L. (2019). La simulación clínica como estrategia para el desarrollo del pensamiento crítico en estudiantes de medicina. *Investigación En Educación Médica*, 8(29), 13–22.
- [Vázquez-Mata, 2008] Vázquez-Mata, G. (2008). Realidad virtual y simulación en el entrenamiento de los estudiantes de medicina. *Educación Médica*, 11, 29–31.
- [Barmaki & Hughes, 2015] Barmaki, R., Hughes, C. E. (2015). Providing real-time feedback for student teachers in a virtual rehearsal environment. *Proceedings of the 2015 ACM on International Conference on Multimodal Interaction*, 531–537.
- [Wyss et al., 2021] Wyss, C., Rosenberger, K., Bühner, W. (2021). Student teachers' and teacher educators' professional vision: Findings from an eye tracking study. *Educational Psychology Review*, 33, 91–107.

- [Goldberg et al., 2021] Goldberg et al., 2021. Goldberg, P., Sümer, Ö., Stürmer, K., Wagner, W., Göllner, R., Gerjets, P., ... Trautwein, U. (2019). Attentive or not? Toward a machine learning approach to assessing students' visible engagement in classroom instruction. *Educational Psychology Review*, 1-23.
- [Cruz et al., 2014] Cruz, J. A. F., Gallardo, P. C., Villarreal, E. A. (2014). La realidad virtual, una tecnología innovadora aplicable al proceso de enseñanza de los estudiantes de ingeniería. *Apertura*, 6(2), 1-10.
- [SÁNCHEZ-CABRERO et al., 2019] SÁNCHEZ-CABRERO, R., NOVILLO-LÓPEZ, M. A., ARIGITA-GARCÍA, A., COSTA-ROMÁN, Ó., BARRIENTOS-FERNÁNDEZ, A., PERICACHO-GÓMEZ, F. J. (2019). Carencias y limitaciones que afectan al asentamiento de la realidad virtual como tecnología de referencia en la sociedad actual. *Revista ESPACIOS*, 40(10).
- [Juncosa, 1999] Juncosa, J. E. (1999). *Etnografía de la comunicación verbal shuar*. Editorial Abya Yala.
- [Cestero Mancera, 2014] Cestero Mancera, A. M. (2014). *Comunicación no verbal y comunicación eficaz*.
- [Rodríguez, 1984] Rodríguez, G. M. (1984). La comunicación humana. Reflexión filosófica en el centenario de Karl Jaspers. *El Guiniguada*, 1, 159-166.
- [Parra, 1988] Parra, M. (1988). La hipótesis sapir-whorf. *Forma y función*, (3), 9-16.
- [Vargas, 2016] Vargas, P. D. P. G. (2016). IMPORTANCIA DE LA COMUNICACIÓN NO VERBAL EN LOS PROCESOS PSICOPEDAGÓGICOS. *Estudios Generales*, 1(1), 8-13.
- [Romo, n.d.] Romo, P. R. (n.d.). *Ensayo sobre*.
- [Hall, Edward. T. , 1959] El lenguaje silencioso. Alianza editorial. Madrid
- [Ponce, 1980] Ponce, F. (1980). BIRDWHISTELL, Ray L.: El lenguaje de la expresión corporal (Book Review). *Arbor*, 105(412), 566.
- [Pereiro, 2019] Pereiro, J. C. (2019). Sin palabras: génesis y desarrollos de los estudios sobre la comunicación no verbal. *Question*.
- [Van der Hofstadt Román, 2005] Van der Hofstadt Román, C. J. (2005). *El libro de las habilidades de comunicación*. Ediciones Díaz de Santos.
- [Orzáiz, O. , 2009] Orzáiz, O. (2009). Comunicación no verbal y paraverbal en el debate político entre Zapatero y Rajoy. *Tonos Digital*, 18. Recuperado de <https://goo.gl/M4mQrn>
- [Ridao Rodrigo, 2017] Ridao Rodrigo, S. (2017). «Es un lector, no un orador»: sobre la tricotomía comunicación verbal, paraverbal y no verbal. *Revista Digital de Investigación en Docencia Universitaria*, 11(1), 77-192.
- [Mehrabian, 1971] Mehrabian, A. (1971). *Silent messages* (Vol. 8, Issue 152). Wadsworth Belmont, CA.

- [Yaffe, 2011] Yaffe, P. (2011). The 7 % rule: fact, fiction, or misunderstanding. *Ubiquity*, 2011(October), 1-5.
- [Ekman & Davidson, 1994] Ekman, P. E., Davidson, R. J. (1994). *The nature of emotion: Fundamental questions*. Oxford University Press.
- [Darwin et al., 1971] Darwin, C., Fuster, J., Oliver, M. A. (1971). *Teoría de la evolución* (Vol. 36). Península.
- [Normoyle et al., 2013] Normoyle, A., Liu, F., Kapadia, M., Badler, N. I., Jörg, S. (2013, August). The effect of posture and dynamics on the perception of emotion. In *Proceedings of the ACM symposium on applied perception* (pp. 91-98).
- [Levav, 2005] Levav, M. (2005). Neuropsicología de la emoción. Particularidades en la infancia. *Revista Argentina de Neuropsicología*, 5, 15-24.
- [Bentinck, 2018] Bentinck, R. (2018). *Estrategias Comprobadas para el Manejo del Aula*. Babelcube Inc..
- [Doyle, 2013] Doyle, W. (2013). Ecological approaches to classroom management. In *Handbook of classroom management* (pp. 107–136). Routledge.
- [Schleicher, 2020] Schleicher, A. (2020). Teaching and Learning International Survey (TALIS) 2018. Insights and Interpretations. OECD.
- [Sarceda-Gorgoso et al., 2020] Sarceda-Gorgoso, M. C., Santos-González, M. C., Rego-Agraso, L. (2020). Las competencias docentes en la formación inicial del profesorado de Educación Secundaria. *Profesorado, Revista de Currículum y Formación Del Profesorado*, 24(3), 401–421.
- [Emmer & Stough, 2001] Emmer, E. T., & Stough, L. M. (2001). Classroom management: A critical part of educational psychology, with implications for teacher education. *Educational Psychologist*, 36(2), 103–112.
- [Mur, 2008] Mur, M. A. (2008). La comunicación no verbal en el aula. *Padres y Maestros/Journal of Parents and Teachers*, 314, 9–13.
- [Gutiérrez Fernández, 2007] Gutiérrez Fernández, J. M. (2007). El comportamiento no verbal en el aula. *Episteme*, 27(1), 167–172.
- [Shablico, 2012] Shablico, S. (2012). La comunicación no verbal en el aula, un análisis en la enseñanza disciplinar. *Cuadernos de Investigación Educativa*, 3(18), 99–121.
- [Lhommet & Marsella, 2014] Lhommet, M., & Marsella, S. C. (2014). Expressing emotion through posture. *The Oxford Handbook of Affective Computing*, 273.
- [Cohn et al., 2007] Cohn, J. F., Ambadar, Z., & Ekman, P. (2007). Observer-based measurement of facial expression with the Facial Action Coding System. *The Handbook of Emotion Elicitation and Assessment*, 1(3), 203–221.
- [Tarnowski et al., 2017] Tarnowski, P., Kołodziej, M., Majkowski, A., & Rak, R. J. (2017). Emotion recognition using facial expressions. *Procedia Computer Science*, 108, 1175–1184.

- [Pato & Guerrero, 2018] Pato, V. M. Á., & Guerrero, R. V. (2018). Aproximación al reconocimiento de emociones faciales basado en posición de puntos de interés. *Pistas Educativas*, 39(128).
- [Iglesias et al., 1989] Iglesias, J., Loeches, A., & Serrano, J. (1989). Expresión facial y reconocimiento de emociones en lactantes. *Infancia y Aprendizaje*, 12(48), 93–113.
- [Sistema de codificación facial] Sistema de codificación facial: <https://es.wikipedia.org/wiki>
- [Wallbott, 1998] Wallbott, H. G. (1998). Bodily expression of emotion. *European Journal of Social Psychology*, 28(6), 879–896.
- [Li et al., 2015] Li, S., Zhu, C., Cui, L., Zhao, N., Li, B., & Zhu, T. (2015). Recognition of emotions using Kinects. *ArXiv Preprint ArXiv:1508.00761*.
- [Macías Gómez & Pinedad Agudelo, 2020] Macías Gómez, M., & Pinedad Agudelo, S. (2020). Generación y control de contenido audiovisual utilizando touchdesigner, ableton live y los movimientos del cuerpo a través de perception neuron (Bachelor's thesis, Escuela de Ingenierías).
- [Kim et al., 2019] Kim, H. S., Hong, N., Kim, M., Yoon, S. G., Yu, H. W., Kong, H.-J., Kim, S.-J., Chai, Y. J., Choi, H. J., & Choi, J. Y. (2019). Application of a perception neuron<sup>®</sup> system in simulation-based surgical training. *Journal of Clinical Medicine*, 8(1), 124.
- [Stavroulia et al., 2019] Stavroulia, K. E., Christofi, M., Baka, E., Michael-Grigoriou, D., Magnenat-Thalmann, N., & Lanitis, A. (2019). Assessing the emotional impact of virtual reality-based teacher training. *The International Journal of Information and Learning Technology*.
- [Riera Sardà, 2012] Riera Sardà, A. (2012). *Computational Intelligence Techniques for Electro-Physiological Data Analysis*.
- [García-Magariño et al., 2019] García-Magariño, I., Cerezo, E., Plaza, I., & Chittaro, L. (2019). A mobile application to report and detect 3D body emotional poses. *Expert Systems with Applications*, 122, 207–216.
- [Moya Hernández & Diosa Silva, n.d.] Moya Hernández, J. S., & Diosa Silva, L. C. (n.d.). Sistema de detección de emociones a partir de la posición corporal utilizando el sensor kinect.
- [Edwards et al., 1998] Edwards, G. J., Cootes, T. F., & Taylor, C. J. (1998). Face recognition using active appearance models. *European Conference on Computer Vision*, 581–595.
- [Uribe 2009] Uribe, L. A. F. (2009). A propósito de la comunicación verbal. *Forma y Función*, 22(2), 121–142.
- [Moré Peláez et al., 2005] Moré Peláez, M. J., Bueno Velazco, C., Rodríguez Atanes, T., & Olivera Zunzunegui, T. (2005). Lenguaje, comunicación y familia. *Humanidades Médicas*, 5(1), 0.

- [Antolinez Cáceres, 1991] Antolinez Cáceres, B. R. (1991). Comunicación familiar. Avances en enfermería.
- [Tarnowski et al., 2020] Tarnowski, P., Kołodziej, M., Majkowski, A., & Rak, R. J. (2020). Eye-Tracking Analysis for Emotion Recognition. Computational Intelligence and Neuroscience, 2020.
- [Montero et al., 2018] Montero, B. M., Cevallos, H. V., & Cuesta, J. D. (2018). Metodologías ágiles frente a las tradicionales en el proceso de desarrollo de software. Espirales Revista Multidisciplinaria de Investigación, 2(17).
- [Suarez et al., 2016] Suarez, L. M. M., Castaño, J. M. S., & Ramos, L. M. J. (2016). Análisis comparativo de las metodologías ágiles en el desarrollo de software aplicadas en Colombia. Gestión Del Talento Humano: Enfoques y Modelos, Corporación Centro Internacional de Marketing Territorial Para La Educación y El Desarrollo-CIMTED, 450–464.
- [Amaro Calderón & Valverde Rebaza, 2007] Amaro Calderón, S. D., & Valverde Rebaza, J. C. (2007). Metodologías ágiles. Perú: Trujillo.
- [Bahit, 2012] Bahit, E. (2012). Scrum y eXtreme Programming para programadores. Autoedición.
- [Muñoz Sanabria, 2013] Muñoz Sanabria, L. F. (2013). Un marco de proceso de software disciplinado basado en la arquitectura y en xp para VSE (Very Small Enterprise) xp/architecture.
- [Duarte & Rojas, 2008] Duarte, A. O., & Rojas, M. (2008). Las metodologías de desarrollo ágil como una oportunidad para la ingeniería del software educativo. Revista Avances En Sistemas e Informática, 5(2), 159–171.
- [García Pino, 2008] García Pino, D. (2008). Adaptación de metodologías de Ingeniería de software orientadas a objeto al mantenimiento evolutivo de aplicaciones. Aplicación a un caso práctico.
- [Castellanos Sánchez & Martínez De la Muela, 2013] Castellanos Sánchez, A., & Martínez De la Muela, A. (2013). Trabajo en equipo con Google Drive en la universidad online. Innovación Educativa (México, DF), 13(63), 75–94.
- [The Eye Tribe] The eye Tribe, la startup danesa que desarrolló Eye Tribe Tracker: [https://en.wikipedia.org/wiki/The\\_Eye\\_Tribe](https://en.wikipedia.org/wiki/The_Eye_Tribe)
- [Cruz-Lara et al., 2013] Cruz-Lara, S., Manjón, B. F., de Carvalho, C. V. (2013). Enfoques innovadores en juegos serios. Ieee Vaep Rita, 1(1), 19–21.
- [Marcano Lárez, 2008] Marcano Lárez, B. E. (2008). Juegos serios y entrenamiento en la sociedad digital.
- [Gallego-Durán et al., 2014] Gallego-Durán, F. J., Villagrà-Arnedo, C.-J., Satorre Cuerda, R., Compañ, P., Molina-Carmona, R., Llorens Largo, F. (2014). Panorámica: serious games, gamification y mucho más.



- [Hwang et al.2013] Yamaguchi K. Kawagoe K. Hwang HH., Ida Y. Development of a virtual classroom for high school teacher training. Lecture Notes in Computer Science, 10011, 2016.
- [Davis, 1989] Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS Quarterly, 319–340.
- [Bocos Corredor et al., 2020] Bocos Corredor, M., López García, Á., & Díaz Nieto, A. (2020). Classroom VR: a VR game to improve communication skills in secondary-school teachers.
- [Schindler et al., 2008] Schindler, K., Van Gool, L., & De Gelder, B. (2008). Recognizing emotions expressed by body pose: A biologically inspired neural model. Neural networks, 21(9), 1238-1246.
- [Inverse Kinematics] Ponderación de vectores de posición basada en Inverse Kinematics: <https://es.mathworks.com>
- [Asset Store] Zona de descarga de paquetes Unity: <https://assetstore.unity.com/>
- [repositorio] Repositorio de GitHub donde podemos encontrar el proyecto de Héctor Martín sobre alarma doméstica basado en Kinect <https://github.com/marcan/kinlarm>
- [Cuestionario post-test] Cuestionario de post-test de la fase de experimentación: <https://docs.google.com/forms/>
- [repositorio de GitHub] Repositorio ubicado en GitHub del proyecto ThreeDPoseUnityBarracuda <https://github.com/>
- [lista de tarjetas gráficas] Listado de tarjetas gráficas que soporta Oculus: <https://support.oculus.com/>
- [Machine Learning Movement] Repositorio de GitHub del proyecto Machine Learning Movement por Kat Sullivan: <https://github.com>
- [Alpha de Cronbach] Método para realizar un Alpha de Cronbach: <https://es.wikipedia.org>
- [AllSky] Asset Store para la descarga de AllSky: <https://assetstore.unity.com/packages/>